Journal of Experimental Psychology

ARTHUR W. MELTON, Editor
Air Force Personnel and Training Research Center

DAVID A. GRANT, Associate Editor
University of Wisconsin

DELOS D. WICKENS, Associate Editor
Ohio State University

CONSULTING EDITORS

E. James Archer, *University of Wisconsin*
Judson S. Brown, *State University of Iowa*
Cletus J. Burke, *Indiana University*
Paul M. Fitts, *Ohio State University*
Frederick C. Frick, *Massachusetts Institute of Technology*
Frank A. Geldard, *University of Virginia*
James J. Gibson, *Cornell University*
Clarence H. Graham, *Columbia University*
Harold W. Hake, *University of Illinois*
Lloyd G. Humphreys, *AF Personnel and Training Research Center*
Arthur L. Irion, *Tulane University*
Howard H. Kendler, *New York University*
Donald B. Lindsley, *University of California, Los Angeles*
Kenneth MacCorquodale, *University of Minnesota*
Quinn McNemar, *Stanford University*
Neal E. Miller, *Yale University*
Kenneth W. Spence, *State University of Iowa*

Arthur C. Hoffman, *Managing Editor*
Helen Ott, *Assistant Managing Editor*
Frances Hanel and Sadie J. Doyle, *Editorial Staff*

CONTENTS

An Information Analysis of Verbal and Motor Responses in a Forced-Paced Serial Task: E. A. Alluisi, P. F. Muller, Jr., and P. M. Fitts 153

Acquisition and Long-Term Retention of a Simple Serial Perceptual-Motor Skill: E. Neumann and R. B. Ammons 159

Antecedents of Sleep: W. B. Webb 162

Stimulus Aftereffects and the Partial-Reinforcement Extinction Effect: S. Katz 167

The Role of Contrast in Stimulus Intensity Dynamism (*V*): K. W. Johnsgard 173

Stimulus Generalization as a Function of Age in Children: S. A. Mednick and L. E. Lehtinen 180

The Properties of Stimuli Associated With Shock Reduction: M. D. Nefzger 184

On the Stimulus and Response in Pursuit Tracking: E. C. Poulton 189

Some Effects of Unequal Spatial Distribution on the Detectability of Radar Targets: P. E. Nicely and G. A. Miller 195

The Role of Class-Descriptive Cues in Paired-Associates Learning: J. J. Wulff and L. M. Stolurow 199

A Deduction from an Excitation-Inhibition Account of Retroactive Inhibition: E. Z. Rothkopf 207

Performance in Eyelid Conditioning Following Interpolated Presentations of the UCS: K. P. Goodrich, L. E. Ross, and A. R. Wagner 214

Estimates of Expected Value as a Function of Distribution Parameters: R. J. Schreiber .. 218

American Psychological Association

Vol. 53 No. 3 March 1957

JOURNAL OF EXPERIMENTAL PSYCHOLOGY

The JOURNAL OF EXPERIMENTAL PSYCHOLOGY is published monthly, two volumes per year, by the American Psychological Association, Inc., at Prince and Lemon Sts., Lancaster, Pa. The subscription rate per volume is $8.00, or $16.00 annually. Single copies are $1.50. Subscriptions, orders, address changes, and business communications should be addressed to the American Psychological Association, Inc., 1333 Sixteenth St. N.W., Washington 6, D. C.

This JOURNAL publishes original experimental investigations which are intended to contribute toward the development of psychology as an experimental science. Studies with normal human subjects are favored over studies involving abnormal or animal subjects, except when the latter are specifically oriented toward the extension of behavior theory. Experimental psychometric studies and studies in applied experimental psychology or engineering psychology may be accepted if they have broad implications for experimental and theoretical psychology.

Normally, articles exceeding 20 printed pages cannot be accepted. Within this limit, the piecemeal, experiment-by-experiment reporting of psychological research is discouraged, and the reporting of the data and conclusions of substantial segments of research effort is encouraged. Specifically, an integrated series of studies accomplished simultaneously (e.g., most doctoral dissertations) *must* be presented in a single article. Special provision is made for the publication of Supplementary Reports and Replication Reports in articles of not more than 1.5 printed pages (see Editor's Note, this JOURNAL, 1957, 53, 1–2).

Address all articles submitted for publication to the editor, Arthur W. Melton, Air Force Personnel and Training Research Center, P. O. Box 1557, Lackland Air Force Base, Texas. All manuscripts should be submitted in two complete copies; only the original copy will be returned if the article is not accepted for publication. Figures may be original line drawings or glossy photographic prints, but should not exceed 8½ by 11 in. Care should be taken to insure legibility of labels and critical features of drawings when reduced to a single-column or (exceptionally) a double-column figure in the JOURNAL.

To save printing costs supplementary material to articles in this JOURNAL are deposited with the American Documentation Institute. Copies may be obtained by ordering the Document Number given in the footnote reference to ADI. Orders should be sent to ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D. C. Orders *must* be accompanied by checks or money orders payable to the Chief, Photoduplication Service, Library of Congress.

Manuscripts submitted to the JOURNAL *must* adhere to the conventions concerning reference citations, preparation of tables and figures, manuscript format, etc. as described in the Publication Manual of the American Psychological Association (*Psychol. Bull.*, 1952, 49 [4, Suppl.], 389–449). In particular, the organization of the manuscript should be text, references, footnotes, tables (one to a page), figure titles (typed on separate page), and figures. All typed material *must* be double-spaced with ample margins. The term "subject," with its various forms, is abbreviated *S*, *Ss*, *S*'s, and *Ss*'; analogous abbreviations are used for "experimenter" (*E*) and "observer" (*O*). When in doubt about the practices of the JOURNAL, authors should examine a recent issue.

Entered as second-class matter, February 6, 1937, at the post office at Lancaster, Pa., under the Act of March 3, 1879. Acceptance for mailing at the special rate of postage provided for in paragraph (d-2), Section 34.40, P. L. & R. of 1948, authorized September 11, 1947.

Send address changes to 1333 Sixteenth St. N.W., Washington 6, D. C. Address changes must reach the Subscription Office by the 10th of the month to take effect the following month. Undelivered copies resulting from address changes will not be replaced; subscribers should notify the post office that they will guarantee second-class forwarding postage. Other claims for undelivered copies must be made within four months of publication.

© 1957 by the American Psychological Association, Inc.

Journal of
Experimental Psychology

VOL. 53, NO. 3 MARCH, 1957

AN INFORMATION ANALYSIS OF VERBAL AND MOTOR RESPONSES IN A FORCED-PACED SERIAL TASK [1]

EARL A. ALLUISI, PAUL F. MULLER, JR.,[2] AND PAUL M. FITTS

The Ohio State University

The results of several recent experiments indicate that choice time (disjunctive reaction time) is approximately a linear function of the average amount of information transmitted per stimulus-response event. This linear relation has been found in a key-pressing task with various numbers of equally probable alternative stimuli up to 10 (**5**), in a lever-positioning task with from two to eight equally likely alternative stimuli (**4**), and in a verbal-response task in which the uncertainty per stimulus was varied in several ways (**6**). Choice time has also been found to decrease linearly as the information lost per response was increased by instructing *S* to respond very quickly so that he made errors (**5**). Although these results were obtained in studies of discrete reactions, similar results have also been obtained with self-paced serial tasks (**1, 3**).

The results of these studies, using either discrete or serial tasks in which *S* determines his own reaction time or response rate, have provided data in support of the original suggestion of Hick (**5**) that the rate of gain of information is a constant. That is, if the amount of information transmitted per response is plotted against the average time per response, a linear function of constant slope is obtained.

The results of studies using forced pacing, in which *E* determines the rate of stimulus presentation and records the frequency and nature of errors, indicate a very different type of function. In this case, the rate of information transmission is plotted against the rate of information presentation. At first very few errors are made and the rate of information transmission increases almost in direct proportion to the rate of information presentation. As the presentation

[1] This research was supported in part by the U. S. Air Force under Contract No. AF 33(616)-43, Project No. 7192, with The Ohio State University Research Foundation, monitored by the Aero Medical Laboratory, and was conducted in the Laboratory of Aviation Psychology. Permission is granted for reproduction, translation, publication, use, and disposal in whole or in part by or for the United States Government. The authors wish to acknowledge the assistance of Hugh B. Martin in the operation of the equipment and of Ilse B. Webb in the analysis of the data, and numerous contributions from Dr. Ralph Queal.

[2] Now at General Electric Advanced Electronics Center, Cornell University, Ithaca, New York.

rate is increased further, however, the rate of information transmission does not keep pace, but reaches a maximum (**10**) and then, either immediately or after some limit is exceeded, it decreases (**7**) as the presentation rate is increased still further.

In commenting on the latter type of function, Quastler (**9**, pp. 343 ff.) has noted that the rate of information transmission may be limited by both the speed of responding and the range of stimulation. The relation is a complicated one, however, because two stimulus parameters determine the rate of information presentation: the rate of occurrence of successive stimuli, and the average uncertainty per stimulus. Thus, the rate of handling information in a forced-paced task may be related in a simple functional way to (*a*) the rate of stimulus presentation, (*b*) the number of alternative stimuli, or (*c*) the joint effect of these two factors expressed as the rate of information presentation. The present study was designed to investigate these relations for both verbal and motor responses to familiar stimuli (Arabic numerals).

Method

The experiment was conducted in two parts which differed only with respect to the type of response required. Ten practiced *S*s responded with motor (key-pressing) responses in Part I, and with verbal (number-naming) responses in Part II, to the forced-paced serial presentation of Arabic numerals. The independent variables were the rate at which the numerals were presented (1, 2, or 3 per sec.) and the number of possible alternative stimuli employed (2, 4, or 8). The over-all effects of learning were minimized by the use of practiced *S*s and familiar Arabic-numeral stimuli. The two types of responses were used in order to lend greater generality to the results and to extend the study of S-R compatibility effects.

Apparatus.—The apparatus used in this experiment was the OSU Serial Discrimeter; this has been described in detail elsewhere (**8**, pp. 12 ff.). The apparatus permitted the use of up to 10 possible stimuli, which could be made to occur in any desired sequence at any given preselected rate. The stimuli were projected onto a 10-in. diameter opal glass screen by means of a 10-unit optical projection system.

The response unit used in Part I consisted of a bank of 10 finger keys. The keys were arranged horizontally in two semicircles to correspond with the natural placement of the 10 finger tips. They were numbered from left to right, *1*, *2*, *3*, . . . , *9*, *0* (*0* represented *10*), and *S* was told to respond to each stimulus numeral by pressing the correspondingly numbered finger key. During Part II, the response unit consisted of a boom microphone connected to a square-wave impulse amplifier, and *S* responded by calling out each successive number into the microphone. Approximately 75 db. of broad-band noise were presented to *S* through earphones in order to mask extraneous sounds during both parts of the experiment.

An automatic scoring unit was used with the apparatus during Part I. With this unit, stimulus and response signals were recorded on a matrix of counters representing the 10 possible stimuli, the 10 possible responses, and a totalizing column. An important aspect of this automatic scoring unit was the time interval during which a response would be scored in association with a previous stimulus. Responses always lag behind stimuli so that under forced-paced conditions a response to one stimulus may be made after that stimulus has disappeared and the next stimulus has been presented. The automatic scoring unit was provided with a "scoring delay" in order to take this lag into account. This could be used to displace the scoring interval by any length of time up to .30 sec. after the presentation of the stimulus. In Part I of the present study, a delay of .25 sec. was used with the two fastest rates of stimulus presentation. This delay period was based on previous work that indicated it to be optimal for these rates of presentation. No delay was needed or used with the slowest rate of stimulus presentation.

The automatic scoring unit was not used during Part II because the apparatus could not discriminate among the different verbal responses. Instead, error patterns were recorded by a monitor using lists of the programmed stimuli. Audio-tape recordings of the verbal responses were taken during Part II and were used to resolve any ambiguity of scoring.

Stimuli.—The stimuli were representations of the Arabic numerals recommended by Baker and Grether (**2**, p. 24). They appeared as .25-in. high light patterns at the center of the display screen, approximately 28 in. in front of

S's eyes. The experimental booth was dimly illuminated, brightness contrast was relatively high, and the stimuli were quite legible.

Subjects.—Ten male *S*s were used in Part I. They were obtained from a pool of students who had volunteered to serve for pay in laboratory experiments. Nine of these *S*s served in Part II. The tenth *S* used in Part II was a member of the laboratory staff who replaced one *S* lost from the original group. This last *S* was given training equivalent to that received by the other *S*s in Part I.

Procedure.—Before Part I of the experiment, *S*s had been pretrained for 16 days (1400 key-pressing responses per *S* per day) with seven different stimulus codes including the specific Arabic numerals used here. In addition to self-paced conditions, this training included five days of forced pacing with rates of stimulus presentation varying from .6 to 1.8 stimuli per sec. After the completion of Part I, but before Part II, all *S*s were pretrained with verbal responses for 12 additional days (700 verbal responses per *S* per day) with the same seven stimulus codes; this training also included five days of forced pacing.

Except for the change from motor to verbal responses, the procedure for both parts of the experiment was identical. Three levels of stimulus complexity (1, 2, and 3 bits/stimulus) were combined with three rates of stimulus presentation (1, 2, and 3 stimuli/sec.) to yield nine conditions for the rate of information presentation. Each *S* responded for one trial at each of these nine combinations. Each trial consisted of the presentation of 100 stimuli, and all nine trials for one *S* occurred during a 1-hr. interval. Each *S* received the three rates of stimulus presentation at a given level of stimulus complexity before he was presented with a different level of complexity. The order of presentation of stimulus complexities and rates was balanced across nine *S*s, and randomly assigned to the tenth.

The 100 stimuli in each trial were random orders (different for each trial) of either 2, 4, or 8 of the Arabic numerals. For the 1-bit (2-alternative) level of complexity there were 50 each of the numerals *4* and *7*. For the 2-bit level there were 25 each of the numerals *3*, *4*, *7*, and *8*. There were either 12 or 13 (randomly decided) of each of the numerals *1* through *0* (which represented *10*), exclusive of *5* and *6*, for the 3-bit complexity level. The fingers used for the key-pressing responses to these three stimulus conditions in Part I were, respectively, the index fingers of each hand, the middle and index fingers of each hand, and all the fingers of each hand exclusive of the thumbs.

Results

Rate of information transmission.—The response data for all 10 *S*s were pooled for the 1000 stimulus presentations at each complexity-rate combination, and the average amount of information transmitted was computed for each condition.[3] This estimate was then divided by the average time for each condition to yield an estimate of the rate of information transmission (H_t, in bits/sec.). The results are shown in Fig. 1 as functions of the rate of information presentation (H_s, in bits/sec.).

There is a large absolute difference between the two response modes in the levels of performance obtained. The maximum rate obtained with verbal responses (7.9 bits/sec.) is about three times as high as the maximum found for the best motor-response condition studied (2.8 bits/sec.).

The general shapes of the functions shown in Fig. 1 are similar for the different numbers of stimuli within each response mode, but differ markedly between verbal and motor responses. There is an increase, on the average, in the rate of information transmission with increases in stimu-

[3] In the computation of information transmission, one response category was used for each alternative stimulus and all omissions were assigned to an additional "no-response" category. Although a failure to respond is certainly qualitatively different from a wrong response, there appears to be no better rational way of handling omissions in computing an over-all estimate of information transmission. In tabulating the omissions in percentage form, an interaction of stimulus presentation rate with stimulus complexity was noted. The percentage of omitted responses was relatively unaffected by increases in stimulus complexity at the lower rates of stimulus presentation, but was markedly increased by increases in stimulus complexity at the higher rates of stimulus presentation. The same sort of interaction was found with the percentage of incorrect responses exclusive of omissions.

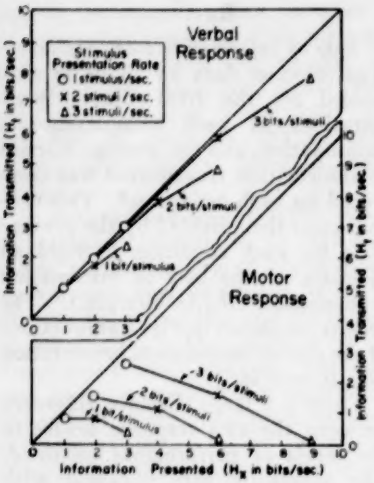

FIG. 1. Absolute rate of information transmission (H_t, in bits/sec.) as a function of the rate of information presentation (H_s, in bits/sec.), for verbal and motor responses at three levels of stimulus complexity and three rates of stimulus presentation. The solid diagonal line represents the maximum possible transmission rate.

lus complexity from 1 to 3 bits/stimulus for both response modes. With increases in the stimulus presentation rate above 1 stimulus/sec., however, there is an increase in the rate of information transmission with verbal responses, but a decrease with motor responses. Thus, for the conditions of this experiment, it is clear that the rate of handling information is not a single function of the rate of information presentation, per se, but is a function of both the number of alternative stimuli and the rate of stimulus presentation. The results further suggest that for both verbal and motor responses, the maximum rate of information transmission will be found when stimulus complexity is extended beyond the level of 3 bits/stimulus used in the present study.

Relative information transmission.—Many tasks place a premium on accuracy and demand that a high proportion of the stimulus information be transmitted, i.e., require a minimum of response equivocation. The present data can also be examined with respect to such a criterion. One measure of the relative efficiency with which information is handled is the ratio of the actual transmission rate to the presentation rate (H_t/H_s, both in bits/sec.). When multiplied by 100, this ratio may be called the *relative information transmission rate*. This is a measure of the percentage of the information associated with the stimulus sequence that can be recovered in the response sequence per unit time. These relative rates, computed from the pooled response data, are shown in Fig. 2 for both verbal and motor responses.

Analyses of variance of these data indicate that there are significant

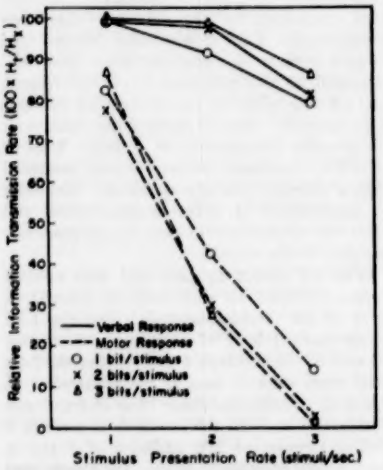

FIG. 2. Relative information transmission rate ($100\ H_t/H_s$, both in bits/sec.) as a function of the rate of stimulus presentation, for verbal and motor responses at three levels of stimulus complexity.

decrements ($P < .01$ in both cases) in the relative information transmission rates for both verbal and motor responses with increases in the rate of stimulus presentation within the range from 1 to 3 stimuli/sec. There are no significant differences ($P > .05$), however, with changes in the number of possible alternative stimuli within the range of stimulus complexities from 1 to 3 bits/stimulus.

Discussion

The large difference between verbal- and motor-response modes could be due in part to practice effects since verbal responses followed motor responses in all cases; this seems unlikely in view of the large number of pretraining responses of both types. On the other hand, the difference may have resulted in part from the slightly different scoring procedures used with the two response modes; there was a greater response-lag tolerance in the nonautomatic scoring of the verbal responses. It is hypothesized, however, that the major portion of the difference in levels of performance should be interpreted as illustrating differences in the degree of S-R compatibility (see **9**, pp. 316–341) between the Arabic-numeral stimuli and the two types of responses. This interpretation is supported by the results of another investigation (**8**) in which the relative superiority of two identical sets of responses was found to be reversed when spatial stimuli (rows of lights) were used in lieu of the Arabic numerals. In other words, it is the combination of stimuli and responses (S-R interaction) that determines the performance level attained rather than the specific set of stimuli or responses per se.

This difference between verbal and motor responses is not as surprising as it might at first appear. If one considers the amount of general experience *Ss* have had in making verbal responses to Arabic numerals as against the experience they have had in making key-pressing responses to the same stimuli, it seems reasonable that the highly overlearned verbal responses might indeed be much more compatible with Arabic-numeral stimuli than are the motor responses.

In each part of the study the information presentation rates of 2, 3, and 6 bits/sec. were produced by two different combinations of stimulus uncertainty and stimulus rate. In the case of each such pair of equal rates of information presentation, the rate of information transmission was greater for the larger number of alternative stimuli and the slower rate of stimulus presentation. This was true for both verbal and motor responses, thus corroborating the findings of Klemmer and Muller (**7**, p. 6) who studied key-pressing responses to flashing light patterns. Within limits, then, for any given rate of information presentation, it appears that an increased rate of information transmission can be obtained by using a larger number of possible alternative stimuli and a slower rate of stimulus presentation. This generalization is further supported by the data on relative information transmission rates.

Summary

The present experiment was designed to determine whether the rate of information transmission in a forced-paced serial task is a function of (*a*) the rate of stimulus presentation, (*b*) the uncertainty per stimulus, or (*c*) the joint effect of these two factors expressed as the rate of information presentation. Arabic numerals were used as stimuli and 10 *Ss* responded to them with both verbal and motor responses. Within the ranges from 1 to 3 bits/stimulus and 1 to 3 stimuli/sec., the data appear to justify the following conclusions.

1. For a given rate of information presentation an increased rate of information transmission was obtained by an increase in the number of possible alternative stimuli and a corresponding decrease in the rate of stimulus presentation. Significant decrements were also found in the relative information transmission rate with increases in the rate of stimulus presentation, but no significant decrements occurred with increases in the number of alternative stimuli. Thus, in forced-paced tasks it

appears that the rate of handling information is not a simple function of the rate of information presentation per se.

2. The differences found between verbal and motor responses to identical stimuli are interpreted as a further instance of variation in S-R compatibility, not as an indication of a general superiority of verbal over motor responses.

3. The results are interpreted as indicating two important interaction effects in forced-paced serial tasks: (*a*) the interaction of stimulus complexity with stimulus rate, and (*b*) the interaction of stimulus code with response mode (S-R compatibility effects).

References

1. Archer, E. J. Identification of visual patterns as a function of information load. *J. exp. Psychol.*, 1954, **48**, 313–317.
2. Baker, C. A., & Grether, W. F. Visual presentation of information. *USAF, WADC Tech. Rep.*, 1954, No. 54-160.
3. Crossman, E. R. F. W. Entropy and choice time: the effect of frequency unbalance on choice-response. *Quart. J. exp. Psychol.*, 1953, **5**, 41–51.
4. Gregg, L. W. The effect of stimulus complexity on discrimination responses. *J. exp. Psychol.*, 1954, **48**, 289–297.
5. Hick, W. E. On the rate of gain of information. *Quart. J. exp. Psychol.*, 1952, **4**, 11–26.
6. Hyman, R. Stimulus information as a determinant of reaction time. *J. exp. Psychol.*, 1953, **45**, 188–196.
7. Klemmer, E. T., & Muller, P. F., Jr. The rate of handling information: key-pressing responses to light patterns. *USAF, ARDC, HFORL Memo. Rep.*, 1953, No. 34.
8. Muller, P. F., Jr. Verbalization as a factor in verbal vs. motor responses to visual stimuli. Unpublished doctor's dissertation, Ohio State Univer., 1955.
9. Quastler, H. (Ed.) *Information theory in psychology: problems and methods.* Glencoe, Ill.: Free Press, 1955.
10. Quastler, H., & Wulff, V. J. Human performance in information transmission. *Univer. Illinois, Control Systems Lab. Rep.*, 1955, No. 62.

(Received January 27, 1956)

Journal of Experimental Psychology
Vol. 53, No. 3, 1957

ACQUISITION AND LONG-TERM RETENTION OF A SIMPLE SERIAL PERCEPTUAL-MOTOR SKILL [1,2]

EVA NEUMANN[3] AND R. B. AMMONS[4]

University of Louisville

Van Dusen and Schlosberg (3) developed a task consisting of serial manipulation of two circles of eight toggle switches, an inner and an outer circle, each switch labeled with a nonsense syllable, in order to study the retention of comparable verbal and motor tasks. In a series of experiments by other investigators (1, 2), it was found that the motor part of the "Schlosberg Board" task was adaptable to the study of perceptual-motor learning per se. It was the purpose of the present study to: (*a*) study the acquisition and retention of this serial perceptual-motor skill, and (*b*) determine the extent to which degree of verbalization used during acquisition has an effect on the serial-position curve.

Method

Twenty male college students served as *Ss* in each of the following five retention groups: 1 min., 20 min., 2 days, 7 weeks, 1 year.

The apparatus has been described in detail elsewhere (3). The *S* sat in front of a large plywood board with the inner and outer circles of switches in "off" position. He turned the inner switches on in clockwise order and attempted to turn on the outer switch that matched each inner switch. Only one switch in the outer circle could be manipulated before going on to the next inner switch. A correct match was signaled by a buzzer. The same pattern of correct matches was used throughout the experiment. The *S* returned each pair of inner and outer switches to the "off" position before going on to the next inner switch. The *S* was paced by a metronome and bell so that he attempted one match each 3 sec. After each complete circuit of the eight inner switches he rested for 1 min., during which time he reported the method he was using. Practice continued until *S* had reached a criterion of two successive errorless trials. Relearning after the retention interval was to the same criterion.

Results and Discussion

The performance curves for acquisition and relearning are shown in Fig. 1. For plotting of the curves, *Ss* were credited with assumed scores of eight correct responses after they had reached the criterion. The points on the curves represent the performance on every third trial; inspection indicated that these are typical of all points. The *N* is 100 at each point for the acquisition curve, and 20 for each of the five retention curves.

It can be seen from Fig. 1 that *Ss* took more trials to relearn the task, the longer the retention interval. Analysis of variance for number of retention trials to criterion showed that differences among retention groups were significant at the 1% level: F was 30.44 with 4 and 95 *df*. During acquisition, the five retention groups did not differ significantly with respect to the number of trials to learn the task to criterion: F was less than 1.00 with 4 and 95 *df*. When *t*

[1] This research was supported in part by the United States Air Force under Contract No. AF 18 (600)-356, monitored by the Air Force Personnel and Training Research Center, Lackland Air Force Base, San Antonio, Texas. Permission is granted for reproduction, translation, publication, use and disposal in whole or in part by or for the United States Government.

[2] The authors gratefully acknowledge helpful suggestions received from Carol H. Ammons and R. Gene Farr, assistance in experimentation by Charles Russell, Criseyde Stone, and Ralph Marion, and help from C. H. Amster in obtaining *Ss*.

[3] Now at Howard University.

[4] Now at University of North Dakota.

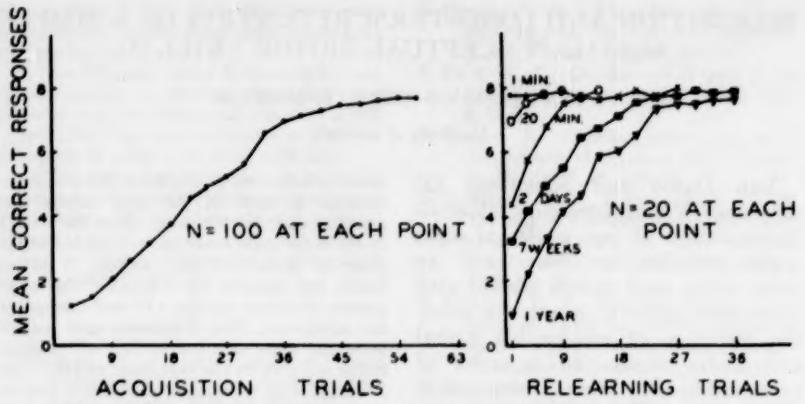

FIG. 1. Performance on every third trial during acquisition and relearning, and on Trial 1 of relearning.

tests were applied to the mean number of relearning trials for the five retention groups, they showed that all differences were significant at the 1% level, except those between the 1-min. and the 20-min. groups and between the 20-min. and the 2-day groups.

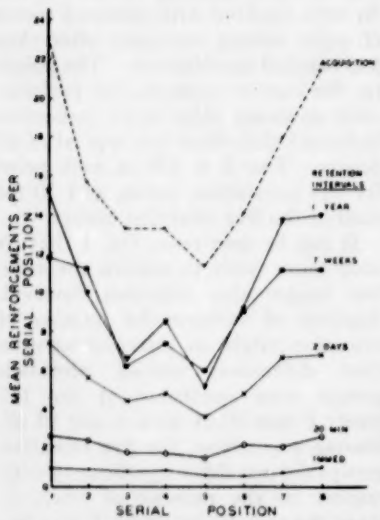

FIG. 2. Serial-position curves for acquisition and relearning periods.

It can also be seen from Fig. 1 that there were fewer correct responses on the first relearning trial, the longer the retention interval. The differences were found by analysis of variance to be significant at the 1% level: F was 6.02 with 4 and 95 df; t tests showed that differences (significant at the 1% level) between the 1-min. and 7-wk. group and 1-yr. group, as well as between the 20-min. group and the 1-yr. group, were primarily responsible for the over-all difference.

Reports by *Ss* of methods used in learning were used to classify *Ss* on the basis of amount of verbalization into the following groups: numbering or counting (44 *Ss*), verbal spatial designation (11 *Ss*), movement patterns (24 *Ss*), visualization (14 *Ss*), and no comment (7 *Ss*). Differences between these groups in mean trials to acquisition criterion did not reach the 10% level of significance. A curve showing mean correct responses at each serial position was plotted for each group. As the performance of every group showed a marked serial-position effect, with poorest performance at the fourth or fifth position,

best at the first and the last, it can be concluded that *serial-position curves are not always dependent on the use of language.*

The serial-position curves for the five retention intervals are shown in Fig. 2. The curves for the 1-min. and 20-min. intervals are approximately straight lines. After a retention interval of 2 days, the serial-position curve has the shape of the curve typical of verbal serial learning; the 7-wk. and 1-yr. curves show an accentuation of this effect; the similarity in pattern between the curves for the three longer retention intervals and the acquisition curve is rather striking.

Summary

This experiment was conducted to determine some characteristics of the acquisition and retention of a simple serial perceptual-motor task. One hundred male college students learned a circular sequence of eight randomly paired toggle switches to a criterion of two consecutive perfect trials. Groups of 20 *Ss* each were retested after 1 min., 20 min., 2 days, 7 weeks, or 1 year, relearning to the same criterion. It was found that: (*a*) *Ss* took longer to relearn the longer the retention interval; (*b*) correct responses on the first relearning trial were fewer the longer the retention interval; (*c*) a marked serial-position effect similar to that observed with serial verbal material appeared during both acquisition and relearning; and (*d*) this serial-position effect was not associated in any detectable way with the degree of verbalization used, suggesting that such effect is due, *not* to the use of verbal material per se, but to the nature of the series learned.

References

1. Ammons, C. H., Slotnick, G., & Ammons, R. B. Performance of a serial paired-associate motor task. *Mot. Skills Res. Exch.*, 1950, **2**, 39–43.
2. Sparks, P. C., & Ammons, R. B. Uses of the Schlosberg Board in the study of human learning. *Mot. Skills Res. Exch.*, 1949, **1**, 21–26.
3. Van Dusen, F., & Schlosberg, H. Further study of the retention of verbal and motor skills. *J. exp. Psychol.*, 1948, **38**, 526–534.

(Received December 21, 1955)

Journal of Experimental Psychology
Vol. 53, No. 3, 1957

ANTECEDENTS OF SLEEP

WILSE B. WEBB[1]

U. S. Naval School of Aviation Medicine[2]

Psychologists have done very little about experimentally analyzing the antecedents of sleep. When they have shown interest in sleep, such interest typically has been directed toward the consequences of sleep deprivation on nonsleep activities or in describing the response itself. Most of the analyses of the causes of sleep have been confined to physiochemical experiments. The studies reported here are concerned with the relationships between the sleep response and the systematic manipulation of three conditions antecedent to that response: past experience, time of sleep deprivation, and an irrelevant drive level.

Experiment I

Experiment I tested the effect of always going to sleep in a particular environment on the rate of going to sleep. A control was introduced to try to distinguish between "habit" and "habituation."

Method

Subjects.—Nineteen female hooded rats were used. At the beginning of the experiment these *S*s were approximately 120 days old. The rats were divided into two weight-matched groups, an experimental group of 10 *S*s (Group A) and a control group of 9 *S*s (Group B).

Sleep measure.—The sleep response was determined by direct observation. After a large number of observations during the familiarization period, consistent agreements were obtained between the two observers based on first a cessation of movement and secondly on an appearance of relaxation of muscle tonus. This judgment was used as a criterion.

Apparatus.—A battery of wakefulness enforcers essentially replicating those developed by Licklider and Bunch (1) was constructed. The apparatus consisted of 24 compartments 4½ by 7½ in. In each compartment was a wheel 4½ in. in diameter which was two-thirds submerged in water. The wheels were driven at a rate of 1 revolution per 108 sec. A food trough was in front of and above each wheel. The compartments were 12½ in. from the top of the wheel to the top of the compartment. The *S*s placed on these moving wheels could readily maintain their position above the water with very minimal movements; however, such movements were required every few seconds.

Observation cages.—Twenty separate wire-mesh observation cages were suspended in a rack. Each cage was 7 by 9½ in., had a 1-w. light at the rear top, and had a smooth bakelite floor. The light and the floor were introduced to attempt to make the observation cages unique from the home cages which were of the same size but were entirely of wire mesh and without a light. The *S*s were housed in groups of three to four, but were observed individually.

Preliminary training.—For a period of five days all *S*s were "familiarized" with the sleep environment and the sleep enforcer apparatus. All *S*s were placed in the observation cages for periods ranging from 10 to 30 min. one or more times per day for a total of 10 hr. per *S* (± 30 min.). In addition, *S*s were placed on the activity wheels for 10 to 90 min. one or more times per day for a total of 20 hr. per *S* (± 20 min.).

Sleep training.—For 17 days the following treatment was given Group A: 0900 hr.—removed from home cage and placed on wakefulness-enforcer wheels; 1400 hr.—removed from wheels and placed in observation cages. The *S*s were permitted to go to sleep and the time was recorded. They were removed from the observation cages after 5 min. of sleep and placed in their home cages. The *S*s showed little disturbance in this transfer.

The following treatment was given Group B: 0800 hr.—*S*s were removed from home cages and placed in observation cages and were kept

[1] The author's appreciation is expressed to Hospital Corpsman J. P. Dippel who aided in the collection of the experimental data.

[2] Opinions and conclusions contained in this report are those of the author. They are not to be construed as necessarily reflecting the view or the endorsement of the Navy Department.

TABLE 1

MEANS, MEDIANS AND RANGES OF DAILY SLEEP LATENCIES (IN MINUTES) FOR THE TRAINING PERIOD OF EXP. I

Day	Mean	Median	Range
1	29.1	24	10–53
2	19.4	18	11–29
3	19.3	17	6–40
4	22.4	23	5–38
5	16.5	14	2–40
6	19.8	18	2–42
7	16.5	18	6–26
8	18.1	17	5–29
9	19.5	18	8–40
10	18.2	17	9–32
11	18.2	18	10–34
12	17.6	18	5–28
13	16.8	16	5–35
14	19.0	21	11–24
15	18.6	17	7–29
16	21.4	16	9–45
17	15.9	17	7–23
All Days	19.1	18	2–53

awake by gently shaking the cage or by poking *S* when he appeared to go to sleep; 0900 hr.—the *S*s were placed on the wakefulness-enforcer wheels; 1400 hr.—the *S*s were removed from the wheels and placed in their home cages. This regimen was followed for Groups A and B except on Days 5, 8, 9, 12, 13, 19, and 20 in a 24-day period.

Results

The mean and median and the ranges of the sleep latencies for Group A are presented in Table 1. An analysis of variance of the 17 days yielded a between-days' variance of 94.37 (16 *df*); a between-*S* variance of 890.36 (9 *df*); and an error variance of 34.95. The between-days' variance was significant at the 1% level. The major effect seems to be in the first day. The first day is significantly different (1% level) from the seventeenth day; the first two days different from the last two days. However, although the average latencies of the last eight days was 18.2 and the first eight days 21.4, this is not a significant difference.

Tests of the sleep "habit."—On Experimental Days 18 and 19, Groups A and B were placed in the observation cages at 0800 hr. and the time to sleep for each *S* was recorded. The two groups were then placed on the activity wheels. At 1400 hr. both groups were removed from the wheels and placed in the observation cages, and the time to sleep for each *S* was recorded. All *S*s were removed from the observation cages and returned to their home cages at the end of 1 hr. The means and *SD*s of individual sums of the latencies (in minutes) for the two days for the two groups before and after the wheel are given in Table 2. An analysis of variance gave no significant effect.

Discussion

There is some evidence to indicate that going to sleep in the same environment results in an increased tendency to go to sleep in an environment as increasing experience with that environment is accrued.

No differences on Days 18 and 19 were obtained between Groups A and B. Although the two groups had been in the observation cages equal numbers of times, Group A had always been permitted to sleep in that environment; Group B had not been permitted to sleep. If the tendency to go to sleep was a function of the activity associated

TABLE 2

MEANS AND *SD*S OF INDIVIDUAL SUMS OF LATENCIES (IN MINUTES) OF TWO TEST DAYS FOR TWO EXPERIMENTAL GROUPS BEFORE AND AFTER WAKEFULNESS-ENFORCER WHEELS

Condition	Group A		Group B	
	Mean	*SD*	Mean	*SD*
Before the wheels	32.70	13.40	28.22	13.56
After the wheels	34.80	17.50	41.89	17.33

with the environment (sleep vs. nonsleep), Groups A and B should have shown differential responses. There was no evidence of such a difference. It would appear then that adaptation to the environment accounted for the increased sleep tendency rather than the development of a sleep habit. It is possible that a more prolonged deprivation period resulting in a greater difference in the sleep behavior of the two groups during the sleep training periods may have resulted in the development of different behavioral responses. However, the present data suggest that the probability of going to sleep is more affected by the repetition of the situation than by the activities associated with the situation.

Experiment II

This experiment analyzed the relation between time of sleep deprivation and latency of going to sleep.

Method

The apparatus and measures used in Exp. I were used in this experiment. The same *Ss* were used but were combined into one group. Experiment II was begun one week after the completion of Exp. I. For varying periods of time prior to 1400, *Ss* were placed on the wakefulness-enforcer wheels. At 1400 they were removed and placed in the observation cages for 1 hr. and the latency of going to sleep was recorded. Recovery days were introduced throughout to minimize the effect of preceding periods of deprivation. The schedule of deprivation and recovery days was as follows: Days 1–2, 30 hr.; Days 3–5, rest; Days 6–7, 20 hr.; Day 8, rest; Day 9, 2½ hr.; Day 10, 0 hr.; Day 11, 0 hr.; Days 12–13, 15 hr.; Days 16–17, 25 hr.; Days 18–20, rest; Day 21, 10 hr.; Day 22, 0 hr.

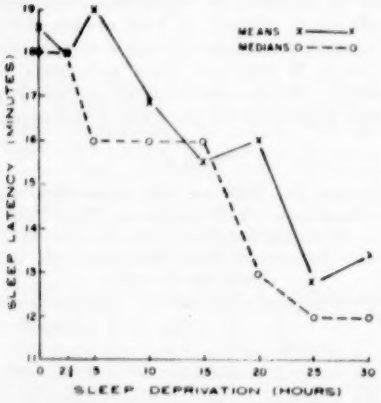

Fig. 1. Sleep latency as a function of the period of sleep deprivation.

Results

Results of the test periods after varying intervals of time are given in Fig. 1. Here is plotted the mean and the median for each test period in terms of the response times (in minutes) in relation to the deprivation intervals. The points plotted for the zero period and the 5-hr. period are the average of three test periods for the zero interval and of two test periods for the 5-hr. interval (Exp. I).

Discussion

There is a relationship between sleep deprivation and time to sleep; however, the data bear discussion.

There was a strong resistance to sleep even after some 30 hr. Thirty hours of sleep deprivation only lowered the latency of sleep by approximately 5 min. when contrasted with limited deprivation. After 30 hr., 3 *Ss* required more than 20 min. to go to sleep, one required 28 min. When *Ss* were taken off the wheel even after long periods of time (25 and 30 hr. of sleep deprivation) they seemed compelled to "adjust themselves" to a nonsleep condition. They went through grooming activity and exploration of their observation cages in spite of the fact that they appeared "exhausted."

It is conceivable that time on the wheel develops a "need to keep awake" which must be extinguished after removal from the wheel. This is the impression that one obtains from ob-

serving *Ss*. If this is true, then the sleep curve which is presented in Fig. 1 represents the requirement for sleep being imposed on top of this need to stay awake rather than being a pure curve of a "need for sleep."

Experiment III

This experiment dealt with the relationship between an irrelevant drive (hunger) and sleep latency.

Method

Again the same apparatus, measures, and *Ss* of Exp. I and II were used. One week after completion of Exp. II food was removed from the home cages of all *Ss* 25 hr. prior to placing them on the wakefulness-enforcer wheels. No food was placed in the feed bins of the wheels. After 5 hr. on the wheel *Ss* were removed and placed in the observation cages for 1 hr. and latencies to sleep were recorded. The observation time was, as in the preceding experiments, at 1400 o'clock. Two observation periods with one week intervening between them were used. The total food deprivation time was 30 hr. at the time of observation. The time to sleep was compared with the two test days of Exp. I which was carried out under 0 hr. of food deprivation and 5 hr. sleep deprivation. The *Ss* on the test days of Exp. I had been fed ad libitum (food in their home cages) up until the time they were placed on the wheels and food was in the food bin of the food compartments.

Results

The group with 5-hr. sleep deprivation and 30-hr. food deprivation had a mean sleep latency of 40.26 min. (SD = 18.63); the group in Exp. I with the same sleep deprivation and no food deprivation required 38.16 min. to sleep (SD = 17.79). The means do not differ significantly.

Discussion

Consistencies in the sleep behavior of individual *Ss* were first noted in Exp. I. An analysis of variance in that experiment yielded a between-*Ss* variance which was proportionately quite large. Rank order correlations between a number of sessions were obtained. In Exp. I the rank order correlation between latencies on Days 1 and 2 was .84, between Days 16 and 17 was .71; the average individual latencies on Days 1 and 2 were correlated .46 with the average latencies of Days 16 and 17. In Exp. II the rank correlation between average individual latencies for the 0-hr. and 30-hr. sleep deprivation was .40. In Exp. III there was a .48 correlation between the average individual latencies of the two 0-hr. food-deprivation tests and the 30-hr. food-deprivation tests.

The range of latencies within an experimental condition was around 30 min. with considerable consistencies in the individual positions within that range. Between experimental conditions the differences between means were more on the order of 5 min. at the outside. It follows then that within the limits of the three variables studied, sleep behavior is far more determined by individual predilections to sleep than by the conditions imposed on the *Ss*.

It was noted in Exp. II that the *Ss* appeared to acquire, at least temporarily, a tendency to stay awake as a result of being on the wakefulness-enforcer wheels. In the extreme periods they seemed to go to sleep *in spite* of this tendency. Even in the shorter periods of wakefulness the *Ss* exhibited tendencies to have to "adjust" to going to sleep. When this is coupled with the knowledge that sleep habits change radically in the direction of increasing periods of wakefulness during the life cycle of human organisms, it is suggested that these experiments may be dealing with the development and existence of wakefulness tendencies as well as the care and nurture of sleep tendencies. Future research will be oriented in that direction.

Summary

In Exp. I there was some evidence of a decrease in sleep latencies in white rats over 17 test days when they were tested in the same environment after 5 hr. of sleep deprivation.

However, the terminal latencies did not differ from the latencies of *Ss* who had had similar treatment but had not been permitted to sleep in the test environment. In Exp. II there was a relationship between sleep deprivation periods of 0, 2½, 10, 15, 20, 25, and 30 hr. and sleep latency, although this effect was not marked. In Exp. III no significant difference was obtained between sleep latencies after 30 hr. of food deprivation versus 0 hr. of food deprivation.

Throughout the experiments a marked individual consistency in latencies of going to sleep was noted across all of the conditions of the experiment. Furthermore, the major determinants of sleep latency were individual differences in contrast to the conditions imposed on these *Ss* (past experience, time of sleep deprivation, and an irrelevant hunger drive). Further observations suggested that the time to sleep may be jointly determined by the development of wakefulness tendencies as well as sleep tendencies.

References

1. Licklider, J. C. R., & Bunch, M. E. Effects of enforced wakefulness upon the growth and the maze-learning performance of white rats. *J. comp. Psychol.*, 1946, **39**, 339–350.

(Received January 23, 1956)

Journal of Experimental Psychology
Vol. 53, No. 3, 1957

STIMULUS AFTEREFFECTS AND THE PARTIAL-REINFORCEMENT EXTINCTION EFFECT[1]

SANFORD KATZ

New York University[2]

The stimulus-generalization hypothesis (**3**) attributes the partial-reinforcement extinction effect to the differential similarity of the stimulus complexes present during partial- and continuous-reinforcement training schedules to the stimulus complex present during extinction. Sheffield (**3**) has attributed this differential similarity to the stimulus aftereffects of reinforcement and nonreinforcement. It is assumed that reinforcement and nonreinforcement result in distinctive stimuli which, if present when the response again occurs and is reinforced, will form part of the conditioned-stimulus complex. When training trials are continuously reinforced, only the aftereffects of reinforcement will serve as part of the CS; when training trials are partially reinforced, both the aftereffects of reinforcement and nonreinforcement will become part of the CS. Since, during extinction, only the aftereffects of nonreinforcement are present, the extinction-stimulus complex will be more similar to the partial-reinforcement CS than to the continuous-reinforcement CS. Hence, on the basis of differential decrement in response strength due to stimulus generalization, partially reinforced *Ss* will demonstrate the greater resistance to extinction.

Sheffield (**3**) tested this hypothesis by comparing the resistance to extinction of partially and continuously reinforced *Ss* when training trials were massed and when they were spaced. On the basis of the assumption that a lengthy intertrial interval would permit the dissipation of the stimulus aftereffects of reinforcement and nonreinforcement, and thus remove the source of the differential similarity, it was predicted that the partial-reinforcement extinction effect would be obtained when training trials were massed but not when they were spaced. This prediction was confirmed.

Recent experiments by Weinstock (**4**) and Katz (**1**) have indicated that the partial-reinforcement extinction effect can be obtained when training trials are spaced. Such demonstrations indicate that the conditioning of the stimulus aftereffects of reinforcement and nonreinforcement, henceforth to be referred to as the *intertrial* factor, is not a necessary condition; other factors, which will be referred to as the *intratrial* factors, may produce the partial-reinforcement extinction effect. However, the results of Sheffield's experiment do suggest that the intertrial factor may be a sufficient condition. The present experiment was an attempt to isolate the intertrial and intratrial factors.

METHOD

Design

Subjects received training in two distinctively different situations, I and II, but extinction trials were only given in Situation II. Each training trial in Situation II was given im-

[1] This article is based upon a portion of a Ph.D. dissertation submitted to the Department of Psychology of New York University. The author wishes to express his appreciation to Dr. M. R. D'Amato, who directed this research.

[2] Now with the U. S. Army Signal Corps Pictorial Center.

mediately after a trial in Situation I, with a lengthy intertrial interval separating consecutive pairs of trials. One group, Group P-C, received a 50%-partial-reinforcement training schedule in Situation I and a continuous-reinforcement training schedule in Situation II. This training procedure provides for the presence of the stimulus aftereffects of reinforcement and nonreinforcement in Situation II. Since these *S*s were continuously reinforced in Situation II, the intratrial factors would not be operative in that situation. A second group, Group C-P, received continuous reinforcement in Situation I and partial reinforcement in Situation II. This training procedure permits the operation of the intratrial factors but not the intertrial factor. The control group, Group C-C, received continuous reinforcement in both situations.

A comparison of the extinction scores of Group P-C with those of Group C-C provides a test of the proposition that the intertrial factor is a sufficient condition for the partial-reinforcement extinction effect. A comparison of Group P-C with Group C-P provides an estimate of the relative importance of the intertrial and intratrial factors in the production of the partial-reinforcement extinction effect.

One factor that would confound the interpretation of the results is the possibility of some form of transfer from Situation I to Situation II. That is, the response strength accruing to the response in Situation I may be transferred to the response in Situation II by virtue of the similarities between the two situations. In addition, the intratrial factors operative in Situation I with Group P-C may generalize to Situation II, thus increasing the resistance to extinction of that group beyond that attributable to the intertrial factor. The amount of transfer can only be reduced by the experimental precautions of making Situations I and II as distinctively different as feasible and by employing a lengthy training series to encourage the formation of a discrimination between the two situations.

Taking into account the possible influence of transfer, and assuming the amount to be, at most, a good deal less than 100%, the following interpretations would be given to the various outcomes:

1. C-P > C-C = P-C would indicate that the conditioning of stimulus aftereffects is not a sufficient condition for the partial-reinforcement extinction effect.

2. C-P > P-C > C-C would indicate that stimulus aftereffects *may* be a sufficient condition, but in any case, not as effective as the intratrial factors.

3. P-C ≧ C-P > C-C would indicate that stimulus aftereffects *are* a factor, and may very well equal or exceed the intratrial factors in importance.

Subjects

The *S*s were male albino rats. Their ages when placed on deprivation were between 71 and 151 days. Seven *S*s were dropped, three for not entering the goal box within 10 min. on three trials during the first day of training, three for failing to eat within 10 min. on one trial during the first day of training, and one for failing to enter the goal box within 5 min. on three trials during the third day of training. Five of these *S*s were dropped from Group C-C, one from Group P-C, and one from Group C-P; leaving Group C-C with 14 *S*s, Group P-C with 16, and Group C-P with 16.

Apparatus

Two straight alleys were employed. The alley used in Situation I was 48 in. long, 6 in. high, and 6 in. wide. It connected a starting box (10 in. long and 6 in. high and wide) to a goal box (12 in. long, 7 in. high, and 6 in. wide). The interior of the alley was painted black, with .5-in.-wide strips of white adhesive tape placed vertically 2 in. apart on the walls throughout the starting box and straight alley. One guillotine door was placed between the starting box and the straight alley, and another was placed in the straight alley 5.25 in. from the goal box. A baffle plate (1.25 in. by 2.5 in.) was placed in the middle of the goal box. Behind it was placed a black food block. Time scores were obtained to the nearest .1 sec. by means of an electric clock. The opening of the starting box door operated a microswitch starting the clock. A 6-in.-long panel, supported by sponge rubber, was situated in the floor 2 in. from the front of the goal box. When the panel was depressed by the weight of *S*, a holding relay was activated, which in turn stopped the clock. A 15-w. frosted bulb was placed 30 in. above and to the side of the goal box.

The straight alley used in Situation II was 49 in. long. It connected a starting box (10 in. long) to a goal box (9 in. long). The whole maze was 4.75 in. wide and 4.75 in. high. The goal box was placed at a right angle to the straight alley. This maze was painted light gray. A guillotine door was placed between the starting box and the straight alley, and another was placed 9 in. from the end of the straight alley. A gray food cup was placed in the far corner of the goal box. A 75-w. frosted bulb was placed 4 ft. above the straight alley, approximately 1

ft. from the starting box. Time scores were obtained in the same way as in Situation I with the 6-in. panel being placed at the front of the goal box.

The *Ss* were housed in a carrying cage during the intertrial interval. The dimensions of each compartment were 8 in. × 3 in. × 5 in.

Procedure

Preliminary training.—The *Ss* were handled frequently prior to the beginning of the experiment. Seven days prior to the beginning of training, *Ss* were put on a 24-hr. feeding schedule, with Purina laboratory chow being available for 1 hr. each day. During this time, *Ss* were given experience eating the 65 mg. food pellets used as reward in the experiment. On the sixth day, each *S* was given 10 min. of free exploration in each of the two situations with the guillotine doors open.

Training.—The 10-day training period began 20 hr. after the end of the feeding period of the sixth day of deprivation. Each *S* was given four trials in each situation on Day 1, six trials on Day 2, and eight trials on each of the remaining eight days. The trials in each alley were given in a single alternation series, with a 20-sec. intertrial interval upon going from Situation I to Situation II, and a 15-min. intertrial interval upon going from Situation II to Situation I.

The order of reinforced and nonreinforced trials for partial-reinforcement training schedules was randomly determined except for the following limitations: (*a*) all trials on Day 1 were reinforced; (*b*) on all other days, 50% of the trials were reinforced; (*c*) the first and last trial of each day were reinforced; (*d*) on Days 2, 3, and 4 no more than two consecutive nonreinforced trials were given.

Sixteen partial-reinforcement orders were chosen from a table of random numbers. No two *Ss* from any one partial reinforcement group received the same order. A reinforced trial differed from a nonreinforced one only in the absence of food reward. On the very first trial in each situation, the food cup was half filled with food pellets, of which *S* was permitted to eat three. On subsequent rewarded trials only two pellets were present. Subjects were left in the goal box for 15 sec., or until the food was consumed, whichever took longer. On nonreinforced trials, *S* remained in the goal box for 15 sec.

Each set of two trials proceeded as follows. The *S* was taken from the carrying cage and placed in the starting box of Situation I with the starting-box door closed and the goal-box door open. After 5 sec. the starting-box door was raised. When *S* stepped on the panel in the goal box, the goal-box door was lowered. After the completion of the trial, *S* was returned to the carrying cage. Twenty seconds later, *S* was placed in Situation II. The procedure in Situation II was the same as in Situation I.

Subjects had food available for 1 hr., 15 min. after the completion of a day's trials. Training on each day began approximately 21 hr. after the end of the last feeding period.

Extinction.—Approximately 21 hr. after the end of the last feeding period, extinction trials were begun. All *Ss* were given 10 extinction trials in Situation II on each of three consecutive days. There was a 20-sec. intertrial interval between each trial. On all extinction trials, *S* was allowed 90 sec. to enter the goal box. If *S* failed to do so, it was removed from the alley and returned to the carrying cage to await the next trial. On completed trials, *S* remained in the goal box for 15 sec. The procedure on any one trial was the same as during training, except for the absence of food.

Results

The response measure employed on each trial was the total time it took *S* to step on the goal-box panel after the starting-box door was opened.

Acquisition.—Figures 1 and 2 present the course of training for each

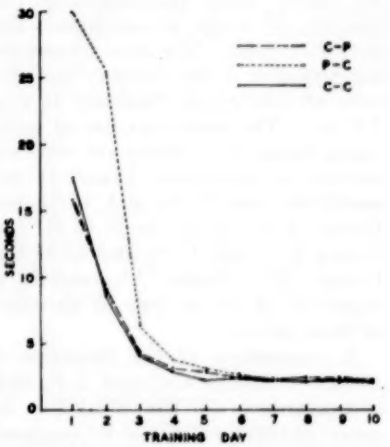

FIG. 1. Acquisition curves in Situation I, showing the average of *Ss'* median total time for each day of training.

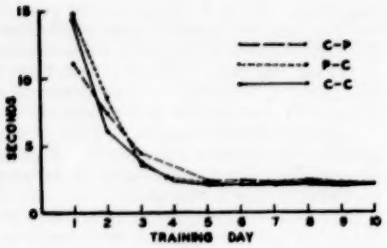

FIG. 2. Acquisition curves in Situation II, showing the average of *Ss'* median total time for each day of training.

day in Situations I and II. Each point on the graphs represents the mean of *Ss'* median total time for each day. It can be seen that, although there were slight differences in favor of the continuous-reinforcement groups during the first five days, all groups had reached the same asymptote by the sixth day. A single measure of each *S*'s performance in each situation on the 30 trials included in the second, third, fourth, and fifth training days was obtained by determining the over-all median for those trials and counting the number of trials at or below the over-all median. The over–all median for Situation I was 3.6 sec. and the over–all median for Situation II was 2.8 sec. The mean number of total times falling at or below the over–all median in Situations I and II respectively was 17.36 and 17.21 for Group C-C, 13.12 and 16.25 for Group P-C, and 15.56 and 12.93 for Group C-P. Table 1 presents a summary of an analysis of variance of these data.

A comparison of the Situation I means of Groups C-C and C-P, and a comparison of the Situation II means of Groups C-C and P-C suggest the presence of transfer from the partial-reinforcement situation to the continuous-reinforcement situation. That is, the factors responsible for the longer total times in the partial-reinforcement situation would appear to have affected *Ss'* performance in the continuous-reinforcement situation. However, the analysis of variance did not indicate a significant effect of percentage of reinforcement. Individual *t* tests of the differences between the Situation I means of Groups C-C and C-P, and of the differences between the Situation II means of Groups C-C and P-C did not result in significant differences. Thus, the differences cannot be taken as proof of the presence of transfer. Conversely, the significant interaction between percentage of reinforcement and situation (see Table 1) indicates that, if there was transfer, the amount was a good deal less than 100%.

Extinction.—The over–all median total time for the 30 extinction trials was 12.6 sec. A score for each *S* was obtained by counting the number of trials which fell at or below this over–all median. The means were 18.25 for Group C-P, 15.06 for Group P-C, and 11.36 for Group C-C. The *F* of an analysis of variance test administered to these data was 11.83 ($P = .001$, $df = 2, 43$). Table 2 presents a summary of the various possible *t* tests. Group C-P demonstrated significantly greater resistance

TABLE 1

SUMMARY OF ANALYSIS OF VARIANCE OF TRAINING RESULTS*

Source of Variation	*df*	*MS*	*F*
% Reinforcement	2	78.82	1.40
Error (b)	43	56.48	
Situation	1	0.39	< 1.00
% Reinforcement × Situation	2	66.50	6.74**
Error (w)	43	9.87	

* See Lindquist (2, pp. 267–273).
** $P = .005$.

TABLE 2

SUMMARY OF t TESTS FOR DIFFERENCES BETWEEN EXTINCTION MEANS

Comparisons (Groups)	Diff.	Error Term	t	P
(C-P) – (C-C)	6.89	1.42	4.85	$.01 > P$
(P-C) – (C-C)	3.70	1.42	2.61	$.05 > P > .01$
(C-P) – (P-C)	3.19	1.37	2.33	$.05 > P > .01$

* Two-tail test, $df = 43$.

to extinction than Groups P-C and C-C; and Group P-C demonstrated significantly greater resistance to extinction than Group C-C. Figure 3 presents the percentage of *Ss* responding at or below the over–all median on groups of five extinction trials. It will be observed that on the third day the three groups were responding at about the same level.

DISCUSSION

The superior resistance to extinction of Group C-P is further evidence that the conditioning of the stimulus aftereffects of previous trials is not a necessary condition for the partial-reinforcement extinction effect. Group C-P had training trials in Situation II with only the aftereffects of reinforcement present during each training trial, and the partial-reinforcement extinction effect was still obtained.

The superior resistance to extinction of Group P-C, as compared with Group C-C, suggests that the conditioning of the stimulus aftereffects of previous trials may be a sufficient condition. However, the possibility of some form of transfer from Situation I to Situation II as is suggested in the training data cannot be ignored. Whether some or all of the superior resistance to extinction could be attributed to the effects of transfer cannot be determined. The same holds true for the importance of the conditioning of the stimulus aftereffects of previous trials.

Although the stimulus aftereffects may be a factor in the partial-reinforcement extinction effect, they cannot be considered as effective as the intratrial factors. The superiority of Group C-P over Group P-C is evidence in favor of this statement. If there were some transfer effects, the differences obtained would underestimate the relative importance of the intratrial factors as compared to the intertrial factor. It therefore seems likely that the intertrial factor is of little importance, and the intratrial factors represent the major factors producing the partial-reinforcement extinction effect.

SUMMARY

The purpose of this study was to compare the effectiveness of stimulus aftereffects and other factors in the production of the partial-reinforcement extinction effect. Rats received training in two distinctively different straight alleys, Situation I and II, with a trial in Situation II given 20 sec. after a trial in Situation I. A 15-min. intertrial interval separated pairs of trials. Group P-C received a partial-reinforcement training schedule in Situation I and a continuous-reinforcement training schedule in Situation II; Group C-P received the reverse of Group P-C; Group C-C received continuous reinforcement in both situations. All *Ss* received 30 extinction trials in Situation II.

The results indicate that Group C-P demonstrated greatest resistance to extinction, with Group P-C next, and Group C-C least. All differences were significant. The presence of some form of transfer from Situation I to

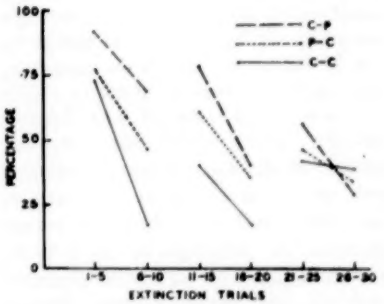

FIG. 3. Extinction curves showing the percentage of responses for consecutive blocks of five trials at or below the over-all extinction median of 12.6 sec.

Situation II is suggested by the training data, and confounds the interpretation of the superiority of Group P-C over Group C-C. The superiority of Group C-P over Group P-C indicates, however, that stimulus aftereffects are of less importance than the other factors.

References

1. Katz, S. An experimental evaluation of the stimulus generalization interpretation of the partial reinforcement extinction effect. Unpublished Ph.D. dissertation, New York Univer., 1955.
2. Lindquist, E. F. *Design and analysis of experiments in psychology and education.* Boston: Houghton Mifflin Co., 1953.
3. Sheffield, V. F. Extinction as a function of partial reinforcement and distribution of practice. *J. exp. Psychol.*, 1949, **39**, 511–525.
4. Weinstock, S. Resistance to extinction of a running response following partial reinforcement under widely spaced trials. *J. comp. physiol. Psychol.*, 1954, **47**, 318–322.

(Received January 20, 1956)

Journal of Experimental Psychology
Vol. 53, No. 3, 1957

THE ROLE OF CONTRAST IN STIMULUS INTENSITY DYNAMISM (*V*) [1]

KEITH W. JOHNSGARD[2]
The State College of Washington

The tendency for stimulus intensity per se to affect response strength has been termed stimulus intensity dynamism by Hull (**8**). He predicts that, other things constant, the magnitude of the stimulus intensity component of reaction potential is a monotonic increasing logarithmic function of stimulus intensity. Since Cattell (**4**) first reported this phenomenon there has been a sizable amount of research which lends support to this formulation. Considering the nature of the relationship exhibited between stimulus intensity and response strength, the majority of investigators have found an increasing negatively accelerated function (**4, 8, 10, 12, 13, 16**), one reported a positive linear relationship (**7**), one reported both positive linear and positive negatively accelerated functions (**2**), one found an **S**-shaped function (**14**), and a few have reported no relationship (**3, 5, 6**). Recent investigators have attempted to separate the effects of stimulus intensity on strength of response (performance) and strength of conditioning (habit), and have uniformly found that stimulus intensity has no significant effect on the strength of conditioning (**5, 6, 10, 11**). It would seem that the dynamism principle is well established.

Recently, however, Bragiel and Perkins (**1, 15**) have introduced stimulus contrast as an alternative to absolute intensity in defining the role of the stimulus in the dynamism concept. These investigators found that a group of rats responding to a white card on a black background in a jumping apparatus exhibited significantly shorter latencies than a similar group responding to a black card on the black background. The results were strikingly similar to those of a like experiment conducted by Hays which Hull reports in support of the dynamism construct (**9**). However, in a second control experiment they found that when a white background was employed, rats responded significantly more rapidly to a black stimulus card. With this one exception, it appears that all of the investigations cited have employed a low intensity background and stimuli which range from a low or medium intensity to a high intensity. Thus, the very nature of these previous experiments would lead to an interpretation of results in terms of absolute physical intensity. While the Bragiel and Perkins study was apparently the first which allowed contrast to operate in both directions, it provided data on only the two extremes of the function. By employing a background of medium intensity with a group of experimental stimuli ranging toward both the low- and high-intensity extremes, it is the purpose of this investigation to help establish whether stimulus intensity

[1] This paper is based on a dissertation presented to the Graduate School of the State College of Washington for the Ph.D. degree. The author expresses his gratitude to Dr. David Ehrenfreund for his advice and aid during the course of the research.

[2] Now at San Jose State College, San Jose, California.

dynamism operates in terms of contrast or in an absolute manner, and to provide data over a wide range of the stimulus continuum.

Method

Design.—Table 1 represents the essential design of this experiment. For simplicity only three experimental stimulus values are included, A background of medium intensity is assumed. Hypothetical values of habit strength ($_sH_R$), generalized habit strength ($_s\dot{H}_R$), and stimulus intensity dynamism (V) contribute to a value in each cell which represents a measure of response strength taken during extinction for the group of *S*s trained and extinguished on that particular stimulus combination.

It will be noted that V operates in an absolute manner (as conventionally defined by Hull) in the top half of Table 1, and operates according to contrast with the medium intensity background in the bottom half. The predicted order of mean response strengths as a function of extinction stimulus intensity (columns) differs markedly according to the assumptions made concerning V.

TABLE 1

Hypothetical Response Strength Values Where V Operates in an Absolute or in a Contrast Manner

Training Stimulus	Extinction Stimulus			Row Means
	Dim	Medium	Bright	
V = Absolute Stimulus Intensity				
Dim	60	64	60	61
Medium	48	80	80	69
Bright	36	64	100	67
Column means	48 <	69 <	80	
V = Stimulus Contrast				
Dim	100	64	60	72
Medium	80	80	80	80
Bright	60	64	100	72
Column means	80 >	69 <	80	

Note.—Where $_sH_R$ = 100, $_s\dot{H}_{R1}$ = 80, $_s\dot{H}_{R2}$ = 60; V = .6, .8, and 1.0 for absolute intensity; and V = 1, .8, and 1 for stimulus contrast.

Subjects.—The *S*s were 50 male and 50 female experimentally naïve albino rats from the colony maintained by the psychology department of the State College of Washington. The age range of *S*s was from 97 to 170 days at the beginning of training.

Apparatus.—The apparatus was a 48-in.-high elevated runway with a starting box at the rear, and a food well depressed below a stimulus card at the anterior end. A large shield was attached to the open end of the starting box preventing *S* from observing *E*. The starting door shortened the box to 8 in. Thus, a 2-in. section of the tunnel extended beyond the door, restricting the *S*'s vision more closely to the area of the experimentally relevant stimuli. A 22½-in.-long and 4-in.-wide runway with ¼-in.-high sheetmetal "guard-rails" extended from the starting door toward the stimulus card. The last 4 in. of the runway consisted of a hinged section resting on a microswitch. The food well was depressed 1½ in. below the end of the runway. A similar distance separated the end of the runway from the stimulus card.

The stimulus cards were rigid with a homogeneous matt surface, and consisted of five 4-in.-wide and 10-in.-high sheets of 060-gauge cellulose acetate which varied along a brightness continuum. Purchased from the Container Corporation of America, these cards had factory calibrated reflectance factors of .035, .13, .34, .57, and .91. When measured by a Macbeth Illuminometer under experimental illumination they had luminances of .09, .54, 1.27, 2.04, and 3.01 ml. When inserted into the end of the apparatus, a 4 × 4-in. section of the card was visible above the runway surface.

A large curved "Cinerama" screen measuring 30 in. high and 40 in. wide was placed 12 in. behind the stimulus card. The matt surface of the screen and the entire apparatus were painted a flat gray which had a luminance of 1.70 ml. under experimental conditions, placing it near the center of the stimulus continuum of experimental cards. Directly behind the screen was a large burlap drape extending from floor to ceiling and curving out to extend 48 in. beyond the sides of the screen. A 150-w. bulb in a glazed glass globe was enclosed in a lightproof box 48 in. above the starting box. A 12-in. circular hole cut in the box directed the light toward the stimulus card and screen.

When *E* opened the starting door, a normally open microswitch was closed, engaging a 110-v. relay and the 110-v. clutch of a Stoelting Chronoscope. When *S* reached the end of the runway, the hinged section dropped a fraction

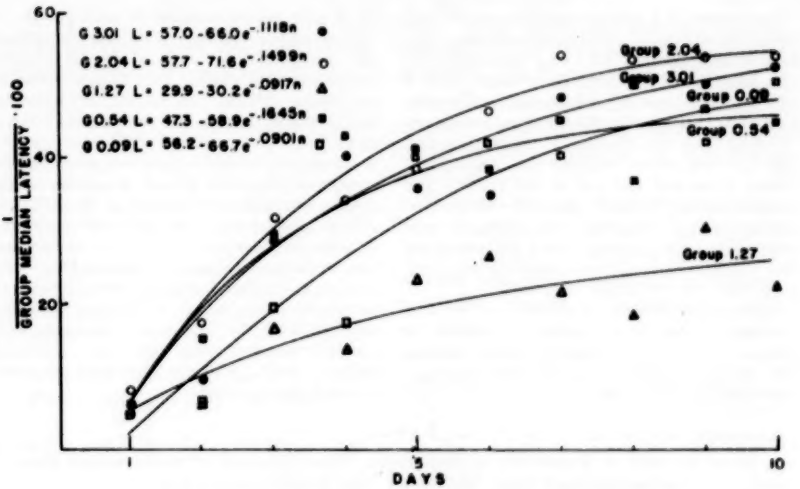

FIG. 1. Acquisition curves of the five groups of 20 *S*s labeled according to the stimulus card luminance to which they responded.

of an inch and opened the normally closed microswitch on which it rested. This opened the relay and stopped the chronoscope.

Preliminary training.—The *S*s were placed on a maintenance schedule 10 days before experience in the apparatus. During this period they were transported to a small feeding room directly adjacent to the experimental room where they were allowed to eat for 1 hr. on individual platforms. The feeding hour was chosen to follow their normal training period by 1 hr. The feeding platforms were 2 in. high, 4 in. wide, and 12 in. long, were extended directly from the wall 48 in. above the floor, and were painted the same medium gray as the apparatus. A glass cup filled with mash was inserted into a hole in the end of the platform most distant from the wall. The mash was 50% water by volume. The solid portion consisted of crushed rat chow. This same chow was used in .2-gm. pellets as reward during the remainder of the experiment. Following the tenth day of maintenance, preliminary training was initiated in the experimental apparatus as follows:

On Day 1 at the regular feeding time *S* was placed on the runway and allowed to feed for 10 min. Reward pellets were scattered over the runway and in the food well. The starting door was closed.

On Day 2 the procedure was identical with the exception that the food was restricted to the food well.

On Day 3 *S* was allowed 5 min. in the apparatus to consume 10 pellets from the food well. The starting door was open and the chronoscope engaged.

On Day 4 *S* was given two distributed trials with the chronoscope engaged.

At no time during preliminary training was a stimulus card present. Further, the hinged section of the runway was always in a locked "up" position.

Training.—Five groups of 20 *S*s were randomly selected to respond to one of the five stimulus cards. The selection was made under the limiting circumstances that 10 males and 10 females be included in each group. Each *S* received 10 trials a day for 10 consecutive days.

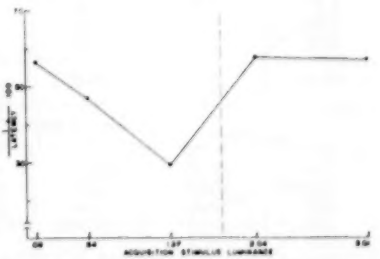

FIG. 2. Acquisition group asymptotes. The values represent the predicted maximas of the five curves indicated in Fig. 1.

A minimum of 5 and a maximum of 15 min. separated the trials. The periods between trials closely approximated the 5-min. minimum with the exception of the intervals between Trials 1 and 2 of any given day which occasionally were as long as 15 min. A typical trial was as follows: When *S* faced the stimulus in the starting box the door was raised engaging the chronoscope. When *S* reached the end of the runway the hinged section dropped, stopping the chronoscope. After consuming the pellet, *S* was returned to the carrying cage. Following the last trial of each day *S*s were allowed to eat mash for 30 min. on the feeding platforms.

Extinction.—The 20 *S*s of each of the five training groups were randomly assigned in groups of four to be extinguished on each of the five stimulus cards. Again the randomization circumstances were limited to insure that two *S*s of each sex be included in each group. A 25-fold matrix resulted, with four *S*s trained and extinguished on each of the 25 possible stimulus combinations. Extinction occurred at the regular running time on the day following the last day of training and proceeded as follows: Each *S* was given one reinforced warm-up trial to its training stimulus, then was given distributed extinction trials (5-min. intervals) to the appropriate extinction stimulus. Each *S* was allowed as many 1-min. periods in the apparatus as were necessary until a criterion of five successive periods without a response was met. During extinction a response was defined as movement to the end of the runway including exploration of the food well. Latencies were recorded for each trial. The cards used for extinction were a duplicate set of those employed in acquisition training.

TABLE 2

MEAN NUMBER OF RESPONSES TO EXTINCTION, MEAN NUMBER OF RESPONSES IN THE FIRST 60 SECONDS OF EXTINCTION, AND MEDIAN LATENCY OF THE FIRST THREE EXTINCTION RESPONSES

Training Stimulus (ml.)	Extinction Stimulus (ml.)					Mean
	.09	.54	1.27	2.04	3.01	
Total Responses						
.09	13.00	11.25	5.75	7.00	11.75	9.75
.54	7.00	8.75	7.50	5.75	10.25	7.85
1.27	8.50	6.25	6.75	8.00	9.00	7.70
2.04	8.75	7.00	5.25	6.25	7.00	6.85
3.01	6.25	7.25	2.75	5.75	7.00	5.80
Mean	8.70	8.10	5.60	6.55	9.00	
Responses in First 60 Sec.						
.09	5.25	4.25	3.75	3.75	7.00	4.80
.54	3.25	2.50	4.00	1.50	2.00	2.65
1.27	3.50	3.00	2.75	5.50	5.00	3.95
2.04	6.25	2.50	4.00	3.50	4.25	4.10
3.01	2.00	3.00	1.50	3.50	2.75	2.55
Mean	4.05	3.05	3.20	3.55	4.20	
Median Latency of First 3 Responses						
.09	47.73	21.71	10.27	22.40	34.00	27.22
.54	10.39	4.19	25.60	2.61	5.73	9.70
1.27	16.61	5.39	12.74	34.99	14.07	16.76
2.04	45.45	26.50	11.65	17.15	30.53	26.26
3.01	9.35	7.06	2.80	37.35	15.57	14.43
Mean	25.91	12.97	12.61	22.90	19.98	

Note—Each of the scores in a given cell is derived from the pooled data of the four *S*s in each cell.

TABLE 3

ANALYSIS OF VARIANCE OF EXTINCTION DATA

Source	*df*	*MS*	*F*
Total Responses			
Extinction stimuli	4	10.65	4.9**
Training stimuli	4	10.62	4.9**
Error	16	2.17	
Responses in First 60 Sec.			
Extinction stimuli	4	1.29	0.9
Training stimuli	4	4.77	3.4*
Error	16	1.41	
Median Latency of First 3 Responses			
Extinction stimuli	4	176.19	1.2
Training stimuli	4	290.66	2.0
Error	16	147.65	

* $P = .05$.
** $P = .01$.

RESULTS

The latency data of the 20 *S*s in each of the acquisition groups were pooled daily and medians computed. The reciprocals × 100 of these median values are plotted in Fig. 1 as a funciton of the day of training. Each empirical point represents 200 trials (10 trials for each of 20 *S*s). The curves were fitted by exponential equations of the general form $L = c - ae^{-id}$, where c is equal to 10, d is the day of training, i is the slope factor, and c is the limiting value of L. These latter c values, which are the predicted maxima of the functions, are plotted in Fig. 2. The dashed line near the center of the continuum represents the background intensity. As a measure of response strength, these values form a roughly U-shaped function indicating that *S*s tended to respond as rapidly to black as to white during training.

The extinction data are included in Table 2, with the three criteria representing measures taken both early and late in the extinction process. The number of responses in 60 sec. was computed by adding the extinction latencies for each *S* until 60 sec. was reached, then simply counting the number of responses included.

F tests were applied to the data. The results are indicated in Table 3. An *F* of 4.9 ($P < .01$) was found when testing the total responses to extinction for variance associated with *extinction* stimulus intensity. Comparable tests of the other two measures failed to approach significance. Response strength means are plotted as a function of extinction stimulus intensity in Fig. 3.

While only the upper curve represents significant variation, the functions are somewhat similar in shape. These roughly U-shaped curves are in line with the predicted columnar mean order indicated in the lower half of Table 1, where *V* functions in terms of contrast. Taken together with the function shown in Fig. 2, they lend support to a contrast interpretation of dynamism.

In testing the total responses to extinction for variance associated with *training* stimulus intensity, an *F* of 4.9 was found ($P < .01$). A similar test of the number of responses in 60 sec. yielded an *F* of 3.4 (*P*

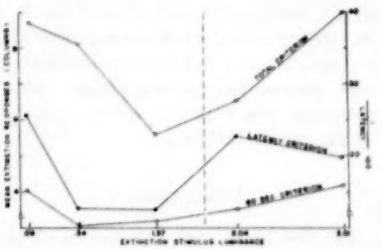

FIG. 3. A plot of the three extinction measures as a function of *extinction* stimulus intensity. The empirical points represent Table 2 columnar means.

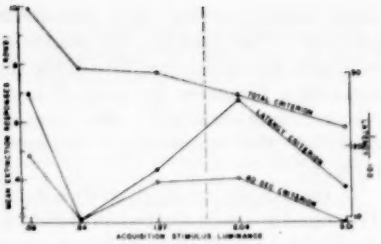

FIG. 4. A plot of the three extinction measures as a function of *training* stimulus intensity. The empirical points represent Table 2 row means.

$< .05$). No significant variation was found when testing the median latency data. Response strength means are plotted as a function of training stimulus intensity in Fig. 4.

The total responses to extinction means, and to a lesser extent the number of responses in 60 sec. means, seem to be related to training stimulus intensity in a meaningful manner, with decreasing stimulus intensity being associated with greater resistance to extinction. The order of these means supports neither of the predictions made in Table 1. Neither do previous investigations support these findings (**5, 6, 10, 11**). It should be noted that there was a greater probability of sampling error in the training stimulus groups than in the extinction stimulus groups. A single randomization placed each *S* in one of the five training stimulus groups, while a second randomization was necessary to determine which stimulus the same *S* would respond to during extinction.

DISCUSSION

For the most part, the results of this investigation point toward a definition of stimulus intensity in terms of *contrast with background intensity* rather than a definition in terms of absolute intensity, when applied to the dynamism concept. In attempting to relate these results to previous investigations and resolve the wide variety of experimental findings, a basic consideration is that in any of the designs the experimental stimulus range could have been restricted to a rather narrow portion of the continuum, either very near or rather distant from the background intensity. It is little wonder then that both linear and curvilinear relationships have been reported, for the investigators might well have been gathering data from either end or the middle of a single large function that could include both linear and curvilinear portions.

On the basis of the Bragiel and Perkins data and the results of the present investigation, let us assume that dynamism operates in terms of stimulus contrast. Since there is an upper limit on response strength, it appears that a theoretical function describing its relationship with degree of stimulus contrast would terminate in a negatively accelerated manner. The majority of experimental findings support such a notion, having been for the most part concerned with experimental stimulus ranges rather removed from the background intensity.

Apparently only one of the experiments cited has been directly concerned with the stimulus range immediately adjacent to the background intensity, and is our only source of experimental information concerning the probable shape of the origin of our theoretical function. The results of that investigation (**14**) show a curve which is first positively, then negatively accelerated. A positively accelerated origin would most probably occur in designs of this nature where the experimental stimulus range is restricted and near the background intensity, and where perception of the experimental stimulus as being different from the background is a prerequisite to making a response. It would appear then, that some sort of an **S**-shaped function describing the relationship between response strength and degree of stimulus contrast with background could adequately account

for the data thus far reported on stimulus intensity dynamism.

Summary

This investigation was designed to study stimulus intensity dynamism (*V*) with special reference to the definition of stimulus intensity in absolute or relative (contrast) terms.

Five groups of 20 rats were trained to run an elevated runway to stimulus cards which varied along a brightness continuum. The stimulus cards were presented before a screen of medium intensity. Each *S* received 100 distributed trials to a given card, during which latencies were recorded. Training groups were then subdivided into groups of four *S*s which were extinguished on one of the five stimulus cards. Extinction data, according to three criteria, were subjected to *F* tests.

For the most part the results tend to support a contrast interpretation of dynamism, and it was suggested that an S-shaped function relating response strength to stimulus contrast with background could adequately account for experimental evidence thus far reported on stimulus intensity dynamism.

References

1. Bragiel, R. M., & Perkins, C. C., Jr. Conditioned stimulus intensity and response speed. *J. exp. Psychol.*, 1954, **47,** 437–441.
2. Brown, J. S. The generalization of approach responses as a function of stimulus intensity and strength of motivation. *J. comp. Psychol.*, 1942, **33,** 209–226.
3. Carter, L. F. Intensity of conditioned stimulus and rate of conditioning. *J. exp. Psychol.*, 1941, **28,** 481–490.
4. Cattell, J. McK. The influence of the intensity of the stimulus on the length of reaction time. *Brain*, 1886, **8,** 512–515.
5. Grant, D. A., & Schneider, D. E. Intensity of the conditioned stimulus and strength of conditioning: I. The conditioned eyelid response to light. *J. exp. Psychol.*, 1948, **38,** 690–696.
6. Grant, D. A., & Schneider, D. E. Intensity of the conditioned stimulus and strength of conditioning: II. The conditioned galvanic skin response to an auditory stimulus. *J. exp. Psychol.*, 1949, **39,** 35–40.
7. Hovland, C. I. The generalization of conditioned responses: II. The sensory generalization of conditioned responses with varying intensities of tone. *J. genet. Psychol.*, 1937, **51,** 279–291.
8. Hull, C. L. Stimulus intensity dynamism (*V*) and stimulus generalization. *Psychol. Rev.*, 1949, **56,** 67–76.
9. Hull, C. L. *Essentials of behavior.* New Haven: Yale Univer. Press, 1951.
10. Kessen, W. Response strength and conditioned stimulus intensity. *J. exp. Psychol.*, 1953, **45,** 82–86.
11. Kimble, G. A. A description of avoidance learning. *Amer. Psychologist*, 1952, **7,** 271. (Abstract)
12. Kupalov, P. S., & Gantt, W. H. The relationship between the strength of the conditioned stimulus and the size of the resulting conditioned reflex. *Brain*, 1927, **50,** 44–52.
13. Kupalov, P. S., Lyman, R. S., & Lukov, B. N. The relation between the intensity of the tone-stimuli and the size of the conditioned reflexes. *Brain*, 1931, **54,** 85–98.
14. Newhall, S. M., & Sears, R. R. Conditioning finger retraction to visual stimuli near the absolute threshold. *Comp. Psychol. Monogr.*, 1933, **9,** No. 43.
15. Perkins, C. C., Jr. The relation between conditioned stimulus intensity and response strength. *J. exp. Psychol.*, 1953, **46,** 225–231.
16. Pieron, H. Nouvelles recherches sur l'analyse du temps de latence sensorielle et sur la loi qui relie le temps à l'intensité d'excitation. *Année Psychol.*, 1920, **22,** 58–142.

(Received February 2, 1956)

Journal of Experimental Psychology
Vol. 53, No. 3, 1957

STIMULUS GENERALIZATION AS A FUNCTION OF AGE IN CHILDREN [1]

SARNOFF A. MEDNICK[2] AND LAURA E. LEHTINEN

Harvard University *Northwestern University*

This study was undertaken to provide data relative to two separate problems. First, the developmental aspects of transposition have become the subject of recent investigations (**1, 6, 10**). In view of the importance of Spence's hypotheses in this area (**9**) and the importance of stimulus generalization (SG) in Spence's hypotheses, a study of the covariability of SG and children's age seemed indicated. A study by Mednick (**7**) has shown that age was not a variable influencing SG in adult males.

Secondly, studies on the generalization of the latencies of a voluntary response have thus far not found latency decreasing regularly as a function of similarity (**4, 5, 8**). Gibson (**5**) has explained this as being due to *E*'s instructions which tend to label generalized responses as errors, causing potential long latency responses to be withheld as incorrect. It was deemed possible that young children would be less facile at inhibiting errors than college students and might yield regular latency generalization gradients.

METHOD

Apparatus.—The apparatus was adapted from one devised by Brown, Bilodeau, and Baron (**4**). It consisted of a horizontal row of 11 lamps fastened to a flat black, curved plywood panel 6 ft. by 2 ft. mounted on its long edge on a table. The lamps were spaced 9° apart and were equidistant from *S*'s eyes when *S* was seated directly in front of and 3.5 ft. away from the center lamp. The lamps will be designated by numbers 1 through 11, with Lamp 1 being on the left of *S*, Lamp 11 on the right of *S*, and Lamp 6 being the center lamp. (The lamps used in this study were 1, 3, 5, 6, 7, 9, 11.) A red, jeweled flashlight lamp 2 in. above the center lamp served as a fixation point and a ready signal. The reaction key was placed on a chair beside *S* on his preferred side. The *E* was seated behind the panel out of *S*'s view. Response frequency and latency were recorded, latency being measured by a Standard Electric Timer.

Subjects.—The *S*s were 72 boys ranging from 7 through 12 yr. The 7- and 8-yr. boys were pupils in Evanston, Illinois, public schools. The *S*s of all other ages were tested at the Evanston YMCA (almost all of these also attended Evanston public schools). Measures of M.A. were unavailable for these *S*s. Six *S*s from the 8-yr. group had to be discarded because of mechanical difficulty with the apparatus (leaving a total *N* of 66). All ages are as of the last birthday.

Procedure.—The *S*s were set to lift their hand from the reaction key as quickly as possible when the center lamp was lit. They were told that other lamps would be lighted occasionally but that they were only to respond to the center lamp. The *S*s were repeatedly encouraged to respond as quickly as possible. In order to provide a behavioral test of *S*'s understanding of the task, two brief demonstration trials were given in which the center lamp and a peripheral lamp were used as test stimuli. It was decided to procure this behavioral indication of *S*'s understanding of the task despite the possibility that the older children would learn more from these trials than the younger children. Ten consecutive training trials with the center lamp (10- to 15-sec. intertrial interval) were then given. The training trials were followed without interruption or warning by a test series during which six of the peripheral lamps (Lamps 1, 3, 5, 7, 9, 11) were presented twice each, interspersed with 17 "booster" trials with the center

[1] We acknowledge the cooperation of Mrs. Dorothy Helding, principal of the Washington School of the Evanston Public Schools, and Mr. Clifford Maxwell of the Evanston YMCA.

[2] This study was conducted while the senior author was on the USPHS Post Doctoral Cooperating Institutions Fellowship. The cooperating institutions are Northwestern University, Neuropsychiatric Institute, University of Illinois, University of Chicago, and Michael Reese Hospital.

TABLE 1

PROPORTION OF *Ss* THAT RESPONDED AT LEAST ONCE TO A GIVEN TEST LAMP

Age Group	*N*	Test Lamps						
		1	3	5	6	7	9	11
7	12	.50	.25	.67	1.00	.75	.67	.58
8	6	.67	.50	.67	1.00	.83	.67	.50
9	12	.75	.58	.67	1.00	.83	.67	.67
Y	30	.63	.43	.67	1.00	.80	.67	.60
10	12	.42	.25	.67	1.00	.67	.42	.25
11	12	.25	.16	.42	1.00	.58	.33	.16
12	12	.33	.25	.50	1.00	.50	.25	.50
O	36	.33	.22	.53	1.00	.58	.33	.30

Note.—Group Y consists of Age Groups 7, 8, and 9. Group O consists of Age Groups 10, 11, and 12.

lamp in a counterbalanced order. The total number of trials in the test series was 29—17 with the center lamp and 12 with the peripheral lamps. Zero, 1, 2, or 3 center-lamp "booster" trials intervened between successive test trials with the peripheral lamps. Six different orders were used for the test trials, each order beginning with a different peripheral lamp. Two *Ss* from each age group were assigned to each order. The *E* promoted faster reaction times by using approximately 50% verbal reinforcement of particularly short reaction times.

RESULTS AND DISCUSSION

The frequency generalization data are presented in Table 1 in the form of the proportion of *Ss* in a group responding to a given lamp. Analysis of variance by ranks was used throughout to evaluate the reliability of differences.

When appropriate, the resulting chi squares were corrected and all were converted into *F* values (11). These are reported in Table 2. As can be seen, the age groups differed significantly in amount of generalization responsivity.

Closer inspection of the data suggested that the differences were, in the main, between the three older age groups (10, 11, 12 years) and the three younger age groups (7, 8, 9 years), each being taken as a composite subgroup. Tests between the age groups within these two composite subgroups revealed no significant differences (Table 2). Thus the three younger groups were combined (Group Y, $N = 30$) and the three older groups were combined (Group O, $N = 36$) into two separate groups. Figure 1 compares the frequency generalization gradients of Group Y and Group O. Both gradients decrease as a function of distance from the center lamp. The trends of the gradients are significant for both Group Y and Group O (Table 2). Both gradients have a sharp drop at Lamp 3 and greater responsivity for Lamps 7, 9, and 11.

TABLE 2

STATISTICAL TESTS OF DIFFERENCES BETWEEN AGE GROUPS AND BETWEEN GENERALIZATION LAMPS

(*F* scores are based on analysis of variance by ranks)

Test	*F*	*df*	*P*
Between all age groups	14.84	5, 23	.01
7 vs. 8 vs. 9 year olds	2.14	2, 8	ns.
10 vs. 11 vs. 12 year olds	3.11	2, 8	ns.
Between generalization lamps (all *Ss*)	9.42	5, 23	.01
Between generalization lamps (Group Y)	8.00	4, 9	.01
Between generalization lamps (Group O)	4.17	4, 9	.05

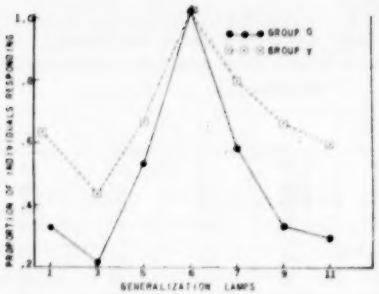

FIG. 1. Generalization gradients based on the proportion of individuals responding at least once at each test lamp in each group.

To test the possibility that differences in the intensity of the lamps may be, in part, responsible for these two aberrancies, measurements of the foot-candle output of the lamps were made. The log foot-candles readings for Lamps 1, 3, 5, 7, 9, 11 were .89, .81, .92, .92, .93, .91, respectively. The difference between Lamp 3 and the other lamps is usually discriminable (**2**). It can be seen that the drop at Lamp 3 and the generally greater responsivity for Lamps 7, 9, and 11 are correlated with similar directional changes in lamp intensity. Since other work has shown covariance of intensity and stimulus generalization responsivity (**3**), it is suggested that such covariance operated in this study. It would seem advisable for future users of this apparatus to counterbalance bulb position or use bulbs equated for foot-candle output.

Reaction times were converted into rates through the use of reciprocal scores. The absence of a response

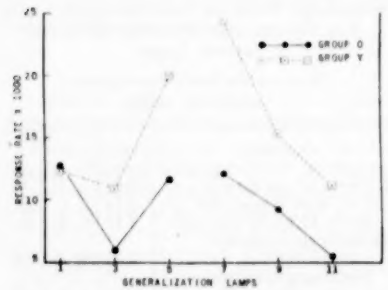

FIG. 2. Generalization gradients based on the mean response rates of each group. An infinite latency was scored as zero.

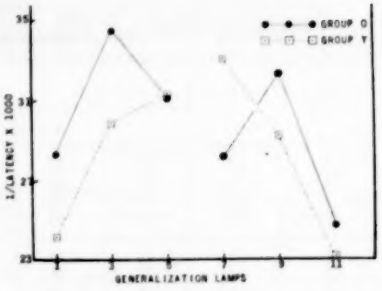

FIG. 3. Generalization gradients based on the reciprocals of mean latencies of each group. Infinite latencies are omitted from consideration.

was then scored as a zero rate. The generalization gradients of response rate are presented in Fig. 2. This gradient closely resembles the form of the frequency gradients even to the drop at Lamp 3. However, closer inspection of the data suggested that the large number of zero rates was a significant cause of the rate gradients mirroring the frequency gradients. Thus the gradients of the reciprocals of the mean latencies (omitting infinite latencies from consideration) are presented in Fig. 3. Group Y demonstrates a regular descending gradient of response rate while Group O has an irregular gradient. In fact, the gradient of Group O resembles the irregular latency gradients found with adult *Ss* (**4, 5, 8**). This finding is seen as lending support to Gibson's explanation of the lack of success in finding latency generalization gradients with adult *Ss*.

It is important to note that while in Fig. 2 Group Y shows an apparently faster response rate than Group O, this difference disappears in Fig. 3. The apparent difference in Fig. 2 is due to the influence of the greater number of zero rates in Group O. To check this, the over-all mean latencies of generalized responses for Groups Y and O were computed and

found to be .3395 and .3337 sec., respectively. A Mann-Whitney *U* test showed these differences to be not significant. This finding negates a possible interpretation which might assert that Group O showed less generalization simply because they were taking more time to respond.

Stevenson, Iscoe, and McConnell (**10**) found that children with a mean age of 7.8 yr. take significantly longer to learn a size discrimination than 11-yr. children. They sensibly interpreted this to be a result of improved speed of learning associated with increased mental age. The results of this study suggest that greater SG reactivity may also in part account for the younger children's slower discrimination learning.

Summary

Measurement of frequency and latency generalization gradients were made for children between the ages of 7 and 12 yr. with an apparatus which measures stimulus generalization along a visual-spatial dimension of similarity and requires a voluntary response. From an analysis of the results, the following conclusions were drawn:

1. Fairly regular decreasing gradients of frequency of response were found for all age ranges.
2. Amount of stimulus generalization reactivity of the younger children (7, 8, 9 yr.) was significantly greater than for the older children (10, 11, 12 yr.).
3. While younger children demonstrate regular increasing generalization gradients of latency, the older children's latency gradients were irregular.

References

1. Alberts, E., & Ehrenfreund, D. Transposition in children as a function of age. *J. exp. Psychol.*, 1951, **41,** 30–38.
2. Boring, E. G. *Sensation and perception in the history of experimental psychology.* New York: Appleton-Century-Crofts, 1942.
3. Brown, J. S. The generalization of approach responses as a function of stimulus intensity and strength of motivation. *J. comp. Psychol.*, 1942, **33,** 209–226.
4. Brown, J. S., Bilodeau, E. A., & Baron, W. R. Bidirectional gradients in the strength of a generalized voluntary response to stimuli on a visual spatial dimension. *J. exp. Psychol.*, 1951, **41,** 52–61.
5. Gibson, E. J. Sensory generalization with voluntary responses. *J. exp. Psychol.*, 1939, **24,** 237–253.
6. Kuenne, M. R. Experimental investigation of the relation of language to transposition behavior in young children. *J. exp. Psychol.*, 1946, **36,** 471–490.
7. Mednick, S. A. Distortions in gradients of stimulus generalization related to cortical brain damage and schizophrenia. *J. abnorm. soc. Psychol.*, 1955, **51,** 536–542.
8. Rosenbaum, G. Stimulus generalization as a function of experimentally induced anxiety. *J. exp. Psychol.*, 1953, **45,** 35–43.
9. Spence, K. W. The nature of discrimination learning in animals. *Psychol. Rev.*, 1936, **43,** 427–449.
10. Stevenson, H. W., Iscoe, I., & McConnell, C. A developmental study of transposition. *J. exp. Psychol.*, 1955, **49,** 278–280.
11. Walker, H. M., & Lev, J. *Statistical inference.* New York: Henry Holt, 1953.

(Received February 6, 1956)

Journal of Experimental Psychology
Vol. 53, No. 3, 1957

THE PROPERTIES OF STIMULI ASSOCIATED WITH SHOCK REDUCTION [1]

M. D. NEFZGER

State University of Iowa [2]

Three recent experiments (**3, 5, 6**) purport to test the hypothesis that stimuli acquire reinforcement value through association with the offset of electric shock. However, the same stimuli could acquire the capacity to elicit a response through such association (**2**); and two of the studies (**3, 6**) do not clearly support the secondary reinforcement hypothesis, despite statistically reliable results, because of testing procedures in which these two acquirable stimulus functions are confounded. There was no such confounding in the third study (**5**), and the hypothesis was not supported.

Two of the studies to be reported involved procedures in which cues like those paired with shock cessation could not have functioned at the choice point of the test maze to elicit responses. The major difference between these studies was in the critical stimulus complex used. In a third study, the critical cues were present in testing at the time of choice and could have elicited an approach response.

METHOD

Experiment I

Subjects.—The *S*s were 10 male and 10 female naive hooded rats of the Iowa Psychology Department strain, 60 to 90 days old when training started, and assigned to two groups at random except that sex and litter were equated.

Apparatus.—A straight alley was used for shock-escape training and a T maze to test for secondary reinforcement. The straight alley was divided by guillotine doors into a starting section (16 in.), runway (16 in.), and endbox (12 in.). All sections were 4 in. wide, 4½ in. deep, and covered by removable Plexiglas. The floor of the starting and runway sections was made of $\frac{3}{32}$-in. brass welding rod placed $\frac{7}{16}$ in. apart. The walls of the starting section and of the 12 in. of the runway adjacent to that section were gray. The walls in the remaining 4 in. of runway will be described along with the endboxes.

Five guillotine doors divided the T maze into a starting section (16 in.), stem (12 in. including the choice area), two arms (12 in.), and two endboxes. The starting, stem, and arm sections were 4 in. wide, 4½ in. deep, and covered by Plexiglas. These sections were gray, and had grid floors the same as those in the straight alley. The two doors in each arm (i.e., one at the choice point and one separating the arm and endbox) were controlled with a single cord so that when one door was lowered the other was raised. Table tops under both alley and maze were gray, and doors were painted the same as sections immediately adjacent.

Two endboxes which provided distinctive visual and tactual cues were used with both pieces of apparatus. The straight-alley endboxes were interchangeable extensions of the runway covered with Plexiglas. Each T-maze endbox consisted of a 12-in. section the same width as the rest of the maze plus a section 12 in. sq.; the narrow part covered with Plexiglas, the square part with ¼-in. mesh hardware cloth. In both cases one of the endboxes was flat black with a wood floor; the other was flat white with fine, white sand in a ¾-in.-deep aluminum tray for the floor.

The 4-in. section of the straight alley adjacent to the endbox was interchangeable and always corresponded in color with the endbox in use. Thus the visual component of the endbox cues occurred some brief time interval prior to escape from the grid floor of the runway, and this condition should enhance the acquisition by those endbox cues of reward properties (**1**).

Shock for escape training was taken from a

[1] This article is based on a disseration submitted to the Graduate College of the State University of Iowa in partial fulfillment of the requirements for the Ph.D. degree. The investigation was carried out under the helpful guidance of Dr. Judson S. Brown.

[2] Now a member of the faculty of the Psychology Department of the University of Buffalo, Buffalo, N. Y.

Variac and through two 6-v. filament transformers wired back to back. Voltage from the functional secondary (117 v.) winding of the second transformer went directly to the parallel wired grids of the straight-alley floor. Shock was never administered in the T maze. Illumination during the experiment was provided by a 200-w., central, overhead source. The two mazes were 180 degrees apart with starting sections toward (and approximately the same distance from) the center of the room.

Procedure.—Twenty days of training were given all *S*s. Eight shock-escape trials were given on Day 1: on Trials 1–4, shock was on as *S*s were dropped into the straight-alley runway; on Trials 5–8, *S*s were put into the starting section through a guillotine door with shock on. Four shock-escape trials were given on Days 2–20: two trials with the shock on as *S*s were put into the straight-alley starting section, the other two with shock off. On the latter type of trial, shock was turned on as soon as possible after *S* was in the start section. On alternate days, either the first and fourth or the second and third trials were given with shock on as *S*s were put into the apparatus. On all trials the door leading to the runway was opened only after *S* had received at least 2 sec. of shock and was oriented toward that door.

The shock was set at 30 v. for Trials 1–4 on Day 1 and 35 v. for Trials 5–8. On Day 2 the shock was 40 v., and thereafter increased in 5-v. steps every other day through Day 10. On Days 11–15 the shock was 65 v., and on Days 16–20 it was 70 v.

One group of 10 *S*s was given shock-escape training using the white endbox, and the other group using the black endbox. After 15–20 sec. in the appropriate endbox, *S* was removed by hand and returned to an individual carrying-cage compartment. The *S*s were run in sequence in squads of five; the sequence varied only by daily rotation of the *S* which was run first.

At least 5 hr. after shock-escape training all *S*s were given four T-maze trials. On Trials 1 and 3, *S* was free to choose either maze arm; on Trials 2 and 4, *S* was forced (by lowering a door at the choice point) into the arm opposite the one chosen on its immediately preceding free trial.[3] On all free trials the doors at the endbox entrances were down when a choice response was made. Only after a complete arm entry was the door into the endbox on that side raised. The corresponding door at the choice point was simultaneously lowered to prevent retracing. In this way it was impossible for endbox cues to elicit approach responses which would directly influence choice behavior; these cues could only reinforce preceding responses.

After endbox entry, the entrance door was lowered; and 30–35 sec. later *S* was removed to the carrying cage. If *S* failed to enter a maze arm within 2 min. after leaving the starting box, or failed to enter the endbox within 2 min. after arm entry, it was removed to the carrying cage. This criterion was invoked on one forced trial.

A random half of the *S*s (three males and two females) in each of the shock-escape training groups was run in the T maze with the white endbox on the left and the black endbox on the right. The other half (two males and three females) was run with the endboxes reversed. The *S*s were run in sequence in squads of 10 with a rotation by days of the *S* to start the sequence.

Experiment II

Twenty additional *S*s identical with those of Exp. I were run under conditions which included a proprioceptive component in the stimulus complex associated with shock cessation. Such stimulation is suggested as critical for the acquisition by that complex of reinforcement value (**6**).

Baffles centered in the interchangeable sections of straight-alley runway and immediately inside the T-maze endboxes forced *S*s to make differential responses. With white endboxes, white horizontal baffles were used: One of each set was near the top of the alley leaving 2½ in. open near the floor; the other was near the bottom leaving 2½ in. at the top. When running into the endbox, *S* had to go first under one then over the next. Vertical, black baffles, also forming a tunnel, were used with black endboxes: One was attached to the left alley wall leaving 1¾ in. open on the right; the other was attached to the right wall leaving 1¾ in. on the left. When running into the endbox, *S* had to go first along the right side of the alley then along the left. Both types of baffles overlapped ¾ in., were 3 in. apart, and were made of ¼-in.-thick aluminum.

Three Microswitches were used to control two Standard Electric time clocks by action of T-maze doors. One timer was activated by lowering the door at the starting-box exit and was stopped by lowering either door at the choice point. This clock provided a measure of time to respond at the choice point. The second timer was activated by lowering either door at the choice point and stopped by raising the same door (which was simultaneous with the

[3] It has been suggested that this procedure gave rise to simple alternation behavior which precluded obtaining results in favor of the hypothesis. There was no evidence from the data to support this suggestion.

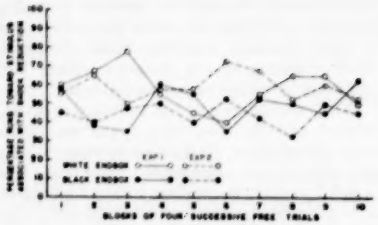

FIG. 1. T-maze choice performance following shock-escape training in Exp. I and II.

lowering of the corresponding endbox door). This clock furnished a measure of time to enter the endbox after a choice of either maze arm.

The first 20 days of the procedure were exactly as in Exp. I. It was necessary to invoke the 2-min. criterion on 24 trials for failure to enter the T-maze endboxes; all occurred before Trial 10. Two additional days of escape training (shock set at 70 v.) were given following the same procedure as on previous days; but in place of T-maze trials, *Ss* were given a simple preference test.

Preference tests were given in a box 20 in. long, 6 in. wide, and 9½ in. deep: One half (10 in. by 6 in.) was white with sand covering the floor, the other half was black with a wood floor. This box pivoted a maximum of $\frac{1}{16}$ in., and a Microswitch activated a Standard Electric time clock when the black side was depressed. The *Ss* were dropped into the middle of the box and removed 10 min. later; time spent in the black side was recorded. The sensitivity of the box was adjusted (by manipulating the quantity and distribution of sand in the white end) following each trial so that a 50-gm. weight placed 2 in. on either side of the pivot tipped the box.

Experiment III

Twenty more *Ss*, 90 to 110 days old at the beginning of training, were run with a procedure which permitted T-maze choice behavior to be determined by the response eliciting capacity of endbox cues. The 12-in. sections of gray alley between the choice point and endboxes were removed from both maze arms, and endboxes were placed next to the choice-point end of the stem. With this modification it was possible to obtain only a single time measure. The 2-min. criterion for response occurrence was invoked on 22 forced and 9 free trials. A forced trial was not given following these 9 free trials. In all other possible respects the procedure of Exp. II was repeated.

RESULTS

If, during shock-escape training, the cues of the escape compartment acquired reinforcing properties, performance in the T maze would reflect an increasing tendency to choose the arm leading to similar cues. The percentage of such choices in successive blocks of four free trials by *Ss* of Exp. I and II is shown in Fig. 1. The comparable data from Exp. III are given in Fig. 2. It is apparent by inspection that systematic increases in percentage of choice were not present in Exp. I and II. The curves in Fig. 2 reflect a modest preference by *both* groups for the black endbox, and this preference appears somewhat stronger on early trials in the group trained with the black escape endbox. The significance of this difference was tested using analysis of variance of a simple mixed (one between and one within subjects component) type; five blocks of eight successive free trials; and number of runs to the black arm per block as the individual measure. No adjustment was made for trials on which *Ss* failed to respond within 2 min. None of the effects (between

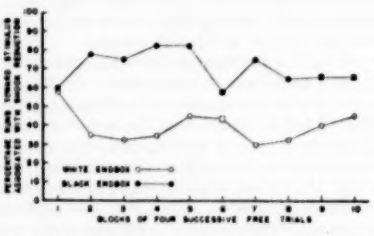

FIG. 2. T-maze choice performance following shock-escape training in Exp. III. The points indicated by squares are based on fewer than 40 trials. The percentage for the sixth block of trials for the group trained with the white shock-escape endbox is based on 37 trials. The three square points for the other groups are each based on 38 trials.

groups, between blocks of trials, and interaction) was significant ($P > .05$).

The evaluation of choice performance using an individual criterion of nine successive turns in the same direction ($P < .05$) (**4**), is given in Table 1. The experimental hypothesis implies that the largest number of *Ss* would reach criterion running to the maze endbox like the one used in shock escape. The statistical hypothesis that it was equally easy to reach criterion running to either endbox (tested using the binomial expansion, $P = .50$, and n = number reaching criterion irrespective of side) could not be rejected at the .05 level in any experiment. There is again in these data the suggestion of a black preference in both groups of Exp. III.

The measures of choice time from Exp. II and III were evaluated in a comparable way by computing individual mean times separately for the last five free and the last five forced runs to each T-maze arm. The means of these individual mean times are given for the two training groups of both experiments in Table 2 along with their estimated standard errors.

Variance analyses of the individual mean times were computed separately for the free and forced trials of each experiment; these were of the same simple mixed type as reported previously. In the four analyses, none of the effects (between training conditions, between maze arms, and their interaction) was significant using the .05 coefficient of risk. Data are excluded for two *Ss* in each group who did not have five free trials to both maze arms.

Time scores obtained from preference testing in the pivoted, black-and-white box were evaluated separately for Exp. II and III using simple mixed analysis of variance. These analyses failed to produce between groups, between days, or interaction effects in either experiment which were significant at the .05 level. In both experiments the mean time spent in

TABLE 1

THE NUMBER OF *Ss* IN ALL GROUPS REACHING THE CRITERION OF NINE SUCCESSIVE TURNS INTO THE SAME T-MAZE ARM

Exp.	Total *N*	Black Shock-Escape Endbox		White Shock-Escape Endbox	
		Black Maze Arm	White Maze Arm	Black Maze Arm	White Maze Arm
I	20	3	3	3	6
II	20	4	4	3	5
III	20	8	1	6	2

TABLE 2

CHOICE TIMES (IN SEC.) ON LAST FIVE FREE AND LAST FIVE FORCED TRIALS TO EACH T-MAZE ARM FOR THE TRAINING GROUPS OF EXP. II AND III

Exp.	*N*	Escape Endbox	Free Trials				Forced Trials			
			Black Maze Arm		White Maze Arm		Black Maze Arm		White Maze Arm	
			Mean	σ_m	Mean	σ_m	Mean	σ_m	Mean	σ_m
II	8	Black	4.73	0.65	5.24	0.85	5.64	1.26	6.10	1.33
	8	White	8.14	1.81	8.20	2.15	7.56	1.69	9.10	1.89
III	8	Black	19.26	5.49	20.34	6.48	23.42	7.00	19.25	7.94
	8	White	25.98	10.30	23.53	6.94	19.63	5.06	17.85	3.66

the black side was largest for the group given shock escape training using the black endbox; both groups spent, on the average, more than half of the time in the black side.

Discussion

There is no evidence in these studies to support the hypothesis that stimuli associated with shock reduction acquire enough reinforcing power to promote learning in a simple-choice situation. Further, there is no clear indication from Exp. III that such stimuli acquire sufficient response-eliciting capacity to determine simple-choice behavior. This latter finding is in opposition to the results of Smith and Buchanan (**6**) and of Goodson and Brownstein (**3**). The former finding is in conformity with that of Littman and Wade (**5**). Comparison of the results of the first two studies fails to support the suggestion of Smith and Buchanan that proprioceptive stimulation is critical to the complex associated with shock offset if that complex is to acquire reinforcing power.

These conclusions must necessarily be tentative because of the large number of procedural differences—other than the conditions of stimulation during test—which could have contributed to the divergence of results. Among these differences is the use by Smith and Buchanan of food reward and its possible interaction with the aftereffects of shock during training and testing.

Summary

Two sets of observations, each involving 20 hooded rats, were taken in an effort to verify the hypothesis that stimuli acquire reinforcement value through association with shock termination. Stimuli like those of a shock-escape compartment were present during testing in one T-maze arm; these stimuli could not have elicited responses at the time of choice. The two studies differed with respect to the stimulus complex used in shock-escape and T-maze endboxes. For a third set of observations on 20 additional *S*s, the critical stimuli could have elicited responses at the T-maze choice point. There was no evidence that stimuli paired with shock offset either reinforce or elicit responses in simple-choice learning.

References

1. Bersh, P. J. The influence of two variables upon the establishment of a secondary reinforcer for operant responses. *J. exp. Psychol.*, 1951, **41**, 62–73.
2. Dinsmoor, J. A. A quantitative comparison of the discriminative and reinforcing functions of a stimulus. *J. exp. Psychol.*, 1950, **40**, 458–472.
3. Goodson, F. E., & Brownstein, A. Secondary reinforcing and motivating properties of stimuli contiguous with shock onset and termination. *J. comp. physiol. Psychol.*, 1955, **48**, 381–386.
4. Hilgard, E. R. Methods and procedures in the study of learning. In S. S. Stevens (Ed.), *Handbook of experimental psychology*. New York: Wiley, 1951. Pp. 517–567.
5. Littman, R. A., & Wade, E. A. A negative test of the drive-reduction hypothesis. *Quart. J. exp. Psychol.*, 1955, **7**, 56–66.
6. Smith, M. P., & Buchanan, G. Acquisition of secondary reward by cues associated with shock reduction. *J. exp. Psychol.*, 1954, **48**, 123–126.

(Received February 16, 1956)

Journal of Experimental Psychology
Vol. 53, No. 3, 1957

ON THE STIMULUS AND RESPONSE IN PURSUIT TRACKING[1]

E. C. POULTON

Applied Psychology Research Unit, Cambridge, England

It was aimed to determine (*a*) what visual information *S* requires in pursuit tracking, and (*b*) the nature of his response with and without this information. The effects of two different kinds of preliminary practice were also compared.

METHOD

Experiment with tachistoscope.—A diagram of the tracking apparatus has been given elsewhere (6). An input (I) pointer moved in a vertical dimension, and *S* had to manipulate a wheel control with a diameter of 6.0 in. to keep a second (R) pointer in line with it. The pointers were triangular in shape. In the present experiment they were a shiny black, and moved against a mat-black background. They were about 1.0 cm. across, and the horizontal gap between their tips was about 5 mm. The largest excursion of the I pointer was 3.8 cm. The wheel control had to be rotated about 30° to match this. A tangential force of about 8.0 oz. was required to overcome the stiction and friction of the control.

The two pointers were viewed with one eye through a box tachistoscope. The box was approximately a cube with sides of 5.0 in. It had a circular opening .5 in. in diameter for *S* to look through. Behind this was a lens which magnified the display and brought it into focus. The box was divided diagonally into two by a thin piece of glass. The *S* either looked straight through the glass at the tracking display, or saw the reflection of a background which was identical with the background of the tracking display. The reflected background could be illuminated by four 3-w. flashlight lamps. The compartment containing the I and R pointers was divided into two by a black vertical partition which separated the two pointers. Each side could be illuminated by two lamps. Thus it was possible for either pointer to be visible all the time, while the other pointer could only be seen intermittently. The experiment was performed in a darkened room, with *S* adequately dark adapted.

Three display conditions were compared. In the intermittent I condition the I pointer was only visible intermittently, while the R pointer could be seen all the time. In the intermittent R condition it was the R pointer which could be seen only intermittently. In the intermittent I and R condition both pointers could be seen only intermittently. With each of these three display conditions, five different lengths of intermittency were used. The intermittent pointer(s) could not be seen for .2, .4, 1.3, 3.0, and 4.7 sec. between glimpses (see Fig. 1). The glimpses lasted about .4 sec. There was also a control period of 60 sec. during which both pointers could be seen all the time, which was followed without a break by a period of 60 sec. during which the intermittent pointer(s) could not be seen at all. The brightness of the glimpses of the intermittent display was matched subjectively with the brightness of the display which could be seen all the time, by reducing the intensity of the latter.

Two different inputs were presented. One was a sine wave with a frequency of 10 cpm. The other input was a combination of three sine waves with frequencies of 5, 7 and 10 cpm. The frequencies had equal amplitudes and random phase relationships. Tracking under each experimental condition lasted 60 sec. A latin-square design was used.

Before the experiment there was a practice on the intermittent I and R condition. The simple input was presented in about 20 periods, each of which lasted 60 sec., and the complex input was presented in the same number of periods. The six *S*s in Group B practiced with brief intervals between glimpses (.05 to .95 sec.). The six *S*s in Group L practiced with longer intervals (.5 to 4.5 sec.). Each *S* was instructed to keep the R pointer in line with the I pointer as well as he could, whether he could see the pointers or not.

Experiment with eyes open and closed.—This

[1] This work was referred to in Applied Psychology Unit Reports 118 and 138 (1950). It was carried out under the general direction of Prof. Sir Frederic Bartlett F.R.S. and Dr. N. H. Mackworth. The experimental *S*s were supplied by the Royal Navy and Royal Air Force. Financial support from the British Medical Research Council is also gratefully acknowledged.

was performed in order to determine whether matching the simple input without seeing the display could be improved by training in the use of kinaesthetic cues. The tracking apparatus was the same as that of the previous experiment, except in the following three respects: (a) *S* used both eyes for viewing the display, as there was no tachistoscope; (*b*) the I and R pointers were white, and the horizontal gap between their tips was only about 2 mm.; and (*c*) the *S* used a horizontal handle, which was fastened to the wheel control and projected 6.0 in. beyond it. An upward movement of the handle raised the R Pointer.

An input of 60 cpm with an amplitude of 5.0 cm. was presented to each of 12 new *S*s. The *S* had to match it for 60 sec. under two conditions. In the experimental condition he had to shut his eyes for alternate periods of 5.0 sec., conforming to *E*'s commands "shut" and "open." In the control condition he had his eyes open all the time. Half the *S*s performed each condition first. A training period of about 30 min. preceded the experiment. During this training *E* frequently described to *S* the nature of the errors which he had made with his eyes shut.

The graphic records (see *Scoring and calculations*) obtained from these 12 *S*s were lost by accident before a detailed analysis of individual responses in the control condition was made. The analysis of the records of six comparable *S*s, who matched a simple input of 60 cpm, has therefore been included instead. These *S*s had had no special training in the use of kinaesthetic cues.

Subjects.—Two of the *S*s in the experiment with the tachistoscope were aircrew in the British Royal Air Force, aged 29 and 30 yr., respectively. The 28 remaining *S*s in the two experiments were enlisted men in the British Royal Navy, aged between 18 and 25 yr. The performance of the two aircrew was no different from that of the remaining *S*s in the same subgroup.

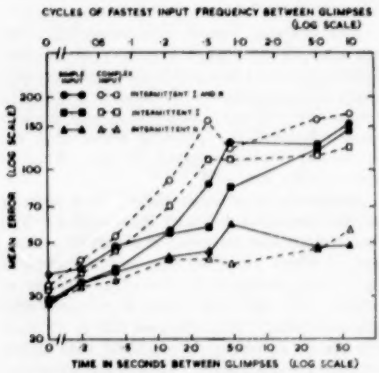

FIG. 1. Mean error with different lengths of interval between glimpses. Each point represents 30 sec. of tracking by each of 12 *S*s. The mean error without tracking was 100.

Scoring and calculations.—Continuous records were obtained on a paper tape of the input, the movement of the R pointer, and of the resultant error. The time of each glimpse was also recorded. The nature of the principal error for each *S* under each condition was determined by inspection of the records. In addition, a score representing the mean error was computed. Units were chosen so that an error of 100 would have been obtained without responding. In the experiment with tachistoscope, the mean error was generally determined for only the second 30 sec. of each experimental period; but when there was no glimpse for 60 sec., the mean error for the first 30 sec. was entered in Fig. 1 as the value for 30 sec. between glimpses, and the mean error for the whole 60 sec. was entered as the value for 60 sec. between glimpses. Simple two-tailed *t* tests were used to determine significance.

RESULTS

Experiment with tachistoscope.—The mean error for the different conditions is shown in Fig. 1. As compared with the continuous view of both pointers, there was a significant deterioration ($P < .05$) in the intermittent R condition with the simple input when the R pointer could not be seen for periods of .2 sec. There was a significant deterioration with the complex input when the R pointer could not be seen for periods of 1.3 sec. The intermittent I condition became significantly less accurate than the intermittent R condition when the respective pointers could not be seen for periods of 4.7 sec. with the simple input, and for periods of .4 sec. with the complex input. The intermittent I and R condition became significantly less accurate than the intermittent I condition when the respective pointers could not be seen for periods of .4 sec.

with the simple input, and for a period of 30 sec. with the complex input.

The main differences in mean error between Groups B and L are shown in Table 1. None of the other measures of performance showed any significant differences between the two groups.

With glimpses separated by intervals of 3.0 and 4.7 sec., the graphic records showed a sudden discrete correction of misalignment following the start of 6% of the glimpses in the intermittent R condition. The corresponding values for the intermittent I and intermittent I and R conditions were 13% and 19%, respectively. The difference between the intermittent R and intermittent I and R conditions was significant at the .001 level. With the remaining glimpses, any correction of misalignment which may have been attempted was combined with matching the movement of the input. The proportion of discrete corrections following glimpses was the same for the 3.0- and the 4.7-sec. intervals between glimpses. There was also no significant difference between the proportions for the two inputs.

When the I pointer could be seen all the time and the R pointer could not be seen at all (intermittent R condition with no glimpses), the absolute position of the response was the most frequent principal source of error, especially with the complex input. The whole response pattern tended to wander slowly upward or downward during the period of 60 sec. With the simple input the amplitude of response was also a fairly frequent principal source of error, the amplitude generally becoming too large.

When the I pointer could not be seen at all (intermittent I and intermittent I and R conditions with no glimpses), the principal source of error was usually in the timing of the response. With the simple input, the error in timing generally took the form of a cumulative error in phase as a result of a small error in wavelength. Responding too early was twice as frequent as responding too late. With the complex input, the timing of the response soon appeared to be unrelated to the input.

TABLE 1

EFFECT OF NATURE OF PRELIMINARY PRACTICE ON MEAN ERROR

Condition	Pooled Mean Error with Intervals Between Glimpses			
	.0 to 1.3 Sec.		3.0 to 60 Sec.	
	Group B	Group L	Group B	Group L
Simple input				
Intermittent R	30	42*	47	54
Intermittent I	34	44*	94	110
Intermittent I and R	38	51*	132	117
Complex input				
Intermittent R	33	36	42	52
Intermittent I	47	46	126†	103‡
Intermittent I and R	59	50	184†	123‡

Note.—The mean error without tracking was 100.
* Pooled Simple input, Group B–Group L $P = .05$.
† Greater than error without tracking ($P = .05$ or better).
‡ Group B–Group L $P = .05$ or better.

Experiment with eyes open and closed.—The mean error for the alternate periods with the eyes closed was 41, as compared with a mean error of 19 in the corresponding control periods. This difference is significant at well beyond the .001 level. However, the mean error without tracking was 100; and before practice, performance with the eyes closed was no better than this.

The distributions of mean errors after practice are shown in Table 2. Eight of the 12 *Ss* tracked as well or better in at least one period with the

TABLE 2

DISTRIBUTIONS OF MEAN ERRORS IN PERIODS OF 5.0 SEC. WITH EYES OPEN AND CLOSED

Condition	Mean Error					
	0–9	10–19	20–29	30–39	40–49	50+
Eyes open	4%	52%	33%	8%	3%	0%
Eyes closed	2%	8%	25%	26%	11%	28%

Note.—The mean error without tracking was 100.

eyes closed, as in one period with the eyes open. About 30% of all attempts with the eyes closed fell into this category. The principal source of error with the eyes closed was generally in the timing of the responses. After each period with the eyes closed, the error in the response movement was rapidly corrected without the R pointer stopping for any appreciable time.

In the control period with the eyes open, the graphic record typically showed areas of several seconds during which the movements of the I and R pointers more or less matched each other. Toward the end of such a period, an error in the phase, amplitude, or absolute position of the movement of the R pointer gradually appeared. This error would then be corrected suddenly, without the movement stopping for any appreciable time. It would be followed by another period of matched I and R movements. For the six additional *S*s, the periods between the discrete corrections varied from 1.0 to 13 sec. The median was 4.0 sec., and about 40% of the periods were 5.0 sec. or longer.

DISCUSSION

Display information required continuously.—Figure 1 shows that in the experiment with tachistoscope, greatest accuracy was achieved when both the pointers could be seen all the time. Even relatively short periods of invisibility of a single pointer increased the mean error. This finding is in line with the results of previously published experiments, in which both pointers were made invisible from time to time (**1**, **3**, **7**).

Figure 1 also shows that *S* could match the simple input of 10 cpm reasonably accurately for periods of up to 3.0 sec., provided he could see one of the two pointers all the time (intermittent I or intermittent **R** condition). Which of the two he saw did not make a significant difference until the other pointer had been invisible for periods of 4.7 sec. By this time an error in the timing of his response was tending to develop in the intermittent I condition. In contrast, with the complex input, reasonable accuracy was possible only in the intermittent **R** condition, for the complex input could not be predicted as successfully by *S* as could the simple input (**7**).

When neither pointer was visible for periods of 3.0 sec. (intermittent I and **R** condition), the mean error with the simple input was not significantly different from the error which would have accrued if *S* had not responded at all. In order to maintain a cyclical movement of about the correct frequency, *S* needed to see a cyclical movement. He had apparently not learned to substitute the kinaesthetic information from his responding arm for this visual information. The experiment with eyes open and closed showed that practice in the use of kinaesthetic cues could improve the matching of a simple input when the display was invisible. This result agrees with that of Lincoln (**4**) for handle winding at a fixed speed.

Intermittency of response.—Craik (**2**) appears to have been the first person to emphasize the intermittent nature of the responses in tracking. He presented relatively unpracticed *S*s with a compensatory display, and noted a discrete response about once every .5 sec. on average (**8**, Fig. 2). He also noted that as *S*s became rather more practiced, these intermittent corrections

of misalignment could have superimposed upon them a constant rate of movement of the control handle (2, p. 59). However, this simple view needs to be extended and modified when detailed studies are made of *S*s responses under more varied conditions.

In the experiment with tachistoscope, only from 6 to 19% of the glimpses separated by 3.0 and 4.7 sec. were accompanied by sudden discrete corrections of misalignment. The misalignment revealed by the remaining glimpses was corrected more gradually, and was combined with an attempt to match the movement of the input. With this pursuit display, discrete corrections of misalignment were thus less evident than with Craik's compensatory display, even though *S* was often presented with a relatively large misalignment during the glimpses.

A dual response, but of a rather different form from that described by Craik, was found in the experiment with eyes open and closed, while tracking the simple input of 60 cpm with eyes open. First, *S* made the R pointer move in much the same way as the I pointer was moving. This required a continuous series of cycles of movement from his responding arm. And secondly, when the movements of the two pointers noticeably failed to match, he adjusted his cyclical response movement to correct the mismatch, without stopping the movement for an appreciable time. For the additional *S*s, these intermittent corrections had a median periodicity of about 4.0 sec.

The dual nature of the response was inferred from the graphic records. However, the finding that about 30% of all attempts with the eyes closed were as accurate as at least one attempt by the same *S* with the eyes open, supports the inference that mismatch with the eyes open was corrected only once every few seconds. The nature of the response in pursuit tracking with high-frequency simple inputs has been described in somewhat similar terms by Noble, Fitts and Warren (5).

The intermittency of the corrections of misalignment may appear to be inconsistent with the finding in the experiment with tachistoscope that tracking was most accurate when both pointers could be seen all the time. But there need be no inconsistency. The former result refers to the output of responses: the response pattern is only modified from time to time to meet the requirements of the incoming signals. While the latter result refers to the incoming signals: the optimal condition in tracking is for signals to be arriving all the time. Presumably, the decision as to when and how mismatch is to be corrected is upset by any loss of visual information. This distinction between the continuous activity of the receptors, and the intermittent modification of responses, does not always seem to have been clearly comprehended (3).

Performance influenced by conditions of preliminary practice.—In the experiment with tachistoscope, Group L did not attempt to match the input as exactly as Group B. Not attempting to match the input exactly was an advantage when the input was complex, and could only be seen relatively infrequently or not at all. Table 1 shows that under these conditions the more conscientious Group B did significantly worse than Group L. Group B actually ran up an error score which was significantly larger than would have accrued if they had not responded at all. But attempting to match the input exactly was the optimal method of performance when the input was simple, and could be seen reasonably frequently or all the time. Under these conditions, Group B was significantly more accurate than Group L.

This difference in performance between the two groups was associated with different conditions of preliminary practice on the intermittent I and R condition. Group L practiced with a range of intervals between glimpses of .5 to 4.5 sec. Under these conditions, not attempting to match the input exactly was an advantage with the complex input. The *S*s presumably noticed this and responded accordingly.

In contrast, Group B practiced with a range of intervals of .05 to .95 sec., where exact matching paid with the simple input. The *S*s were presumably aware of this, and so persevered in their attempts. Thus the method of performance learned in the practice appears to have carried over into the experiment.

Summary

One experiment on pursuit tracking in one dimension used a tachistoscope. In different conditions (*a*) the input (I) pointer, (*b*) the response (R) pointer, (*c*) both pointers, and (*d*) neither pointer could be seen only intermittently, while the other pointer(s) could be seen all the time. Comparisons were made between simple and complex sinusoidal inputs, and also between two different kinds of preliminary practice. A second experiment used a fast simple input. Tracking for alternate periods with the eyes open and closed was compared with normal tracking.

Tracking was found to be most accurate when both pointers could be seen all the time. It was less accurate without seeing the I pointer than without seeing the R pointer, and least accurate when neither pointer could be seen. Sudden discrete corrections of misalignment followed only a minority of the glimpses of the invisible pointer(s).

Tracking for 5.0 sec. with the eyes closed was sometimes as accurate as normal tracking with the eyes open. The fast simple input gave rise to a continuous cyclical response, which appeared to be adjusted only relatively infrequently to improve its match with the input.

References

1. Battig, W. F., Voss, J. F., & Brogden, W. J. Effect of frequency of target intermittence upon tracking. *J. exp. Psychol.*, 1955, **49**, 244–248.
2. Craik, K. J. W. Theory of the human operator in control systems. I. The operator as an engineering system. *Brit. J. Psychol.*, 1947, **38**, 56–61.
3. Katz, M. S., & Spragg, S. D. S. Tracking performance as a function of frequency of course illumination. *J. Psychol.*, 1955, **40**, 181–191.
4. Lincoln, R. S. Learning a rate of movement. *J. exp. Psychol.*, 1954, **47**, 465–470.
5. Noble, M., Fitts, P. M., & Warren, C. E. The frequency response of skilled subjects in a pursuit tracking task. *J. exp. Psychol.*, 1955, **49**, 249–256.
6. Poulton, E. C. Perceptual anticipation in tracking with two-pointer and one-pointer displays. *Brit. J. Psychol.*, 1952, **43**, 222–229.
7. Poulton, E. C. The basis of perceptual anticipation in tracking. *Brit. J. Psychol.*, 1952, **43**, 295–302.
8. Vince, M. A. The intermittency of control movements and the psychological refractory period. *Brit. J. Psychol.*, 1948, **38**, 149–157.

(Received February 7, 1956)

Journal of Experimental Psychology
Vol. 53, No. 3, 1957

SOME EFFECTS OF UNEQUAL SPATIAL DISTRIBUTION ON THE DETECTABILITY OF RADAR TARGETS [1]

PATRICIA E. NICELY

Lincoln Laboratory, Massachusetts Institute of Technology

AND

GEORGE A. MILLER

Harvard University

This paper is concerned with the decline in a radar operator's performance that is frequently observed during a prolonged watch. The literature on this problem has been considered in a theoretical paper by Deese (1). The present concern is to determine (*a*) whether the decline is critically dependent upon the rather difficult viewing conditions imposed by the standard design of radar display units, and (*b*) whether the decline is due to a general effect of relaxation, fatigue, etc., or is a selective phenomenon which can appear in one part of the task and not in another. In order to study (*a*), two different techniques have been used to display the radar data. In order to study (*b*), an unequal spatial distribution of targets has been imposed over the display surface.

Apparatus and Procedure

Isolated target pips were introduced on a simulated radar display[2] that was speckled with a background of random noise. The pip appeared for only one rotation of the antenna. The *S*s reported when they detected a target by calling out its position in terms of numbers on a clock face. In one quadrant of the display, targets appeared on the average of once in every five rotations of the antenna. On the rest of the display the targets averaged only one in every 30 rotations. The *S*s worked for 90 min. and a continuous record was kept of all their reports.

The following procedure was used to determine the levels of the target and the noise. First, the brightness of the strobe line was adjusted to be just visible as it rotated over the face of the CRT, then noise was introduced until it caused a just detectable brightening of the strobe line. The noise was then increased 15 db above this threshold value. Finally, the intensity of the target was varied in a series of preliminary experiments until the level that would yield detections on approximately 60% of the targets was established. Once these levels had been set, they were measured on an oscilloscope and a voltmeter and the measured values were duplicated for each *S*.[3]

The antenna rotated at 6 rpm. The target pip was 2.5° long and 1.9 mm. wide. The noise would best be described as "finely stippled"; the irregular and random nature of the pattern was clearly detectable, but none of the noise "clumps" appeared as large as the target. The effects of changing the grain of the noise (by varying the range on the PPI) were considered, but the impression obtained was that the effects which were to be studied would not be greatly affected by this variable. The experiments were designed to study relative changes in detecta-

[1] The authors wish to express their appreciation to Dr. Herbert M. Jenkins for theoretical advice and criticism and to Dr. M. M. Kessler and B. V. Eckstein for providing and maintaining the electronic equipment used in this study. The research in this document was supported jointly by the Army, Navy, and Air Force under contract with the Massachusetts Institute of Technology.

[2] The radar display was the type known as a plan position indicator (PPI). This means that the relative position of the targets is indicated by the position of a spot of light (the target pip) on the circular, phosphorescent face of a cathode ray tube (CRT). A bright radial line (the strobe line) sweeps around the face of the tube, indicating the rotation of the radar antenna, and "paints in" the target pip at the appropriate range and azimuth.

[3] The electronic details and measurements will be furnished on request to anyone interested in duplicating them.

bility as a function of location and time during the watch, and not to obtain an absolute measure of the signal-to-noise ratio at threshold.

Ten *Ss* were given a short practice session before the test began. They were shown what a target looked like and they demonstrated that they knew what to do by detecting some practice targets. They were advised to let their eyes follow a little behind the strobe line as they searched, since that was where the targets would be the brightest before they faded out. The men were all enlisted Air Force personnel with normal visual acuity, but they had had little or no previous experience as radar observers.

While the tests were being run on the PPI display, a camera was simultaneously taking a picture on each rotation of the antenna. This film was later developed, and when the negative was projected, it provided a bright, steady display of the radar data on each successive rotation. Four of the films suffered mistakes in photography or development and had to be discarded, but the remaining six films were used to test 10 more *Ss* in a manner as similar as possible to the tests that had previously been conducted with the first 10 *Ss* on the CRT display. Thus, there were two separate experiments on the effects of an unequal distribution of targets, one conducted with the standard kind of CRT display and the other with a photographic presentation of the radar data.

Results

The total number of targets and the total number of detections for all 20 *Ss* on both displays in the high- and in the low-probability areas are summarized in Table 1. When all data are combined it is found that 1,524 detections were made of the 2,150 targets that appeared in the high-probability area, and that 222 detections were made of the 354 targets that appeared in the low-probability area, which correspond to detection ratios of .71 and .63, respectively. This difference is significant ($P < .002$), and indicates that *Ss* tended to neglect the targets in the low-probability areas.

Table 1 reveals that more targets were detected with the photographic display than with the CRT display. However, the frequency of false reports was also much higher on the photographic display, with the result that the information was actually less reliable for the photographic display than for the CRT display. Since the performance observed for the film and the CRT is probably a function of the particular noise conditions and the particular photographic processes employed, there is no confidence that the absolute differences observed between the two displays can be generalized beyond the present experimental conditions.

A more interesting feature of the data is the relative decline in *Ss'* performance on low-probability tar-

TABLE 1

Total Number of Targets and Detections as a Function of Time for High- and Low-Probability Areas on Both Displays

Time in Minutes	CRT Display				Photographic Display			
	High Probability		Low Probability		High Probability		Low Probability	
	Targets	Detects.	Targets	Detects.	Targets	Detects.	Targets	Detects.
0–5	57	23	10	7	57	43	10	8
5–15	116	66	20	13	118	97	20	15
15–30	179	117	29	14	180	132	25	18
30–60	363	230	60	33	375	306	60	41
60–90	355	234	60	34	350	276	60	39
Totals	1070	670	179	101	1080	854	175	121

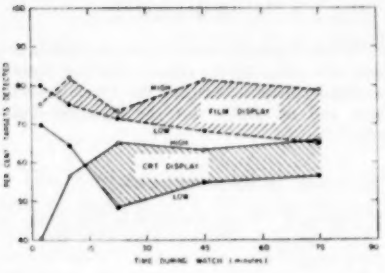

FIG. 1. Per cent targets detected on the CRT and film displays as a function of time during the watch (10 *S*s for each display). By the end of the first 30 min., detection was higher in the area where the probability of a target was higher.

gets during the first 15 min. of the watch. This decline is not immediately apparent from the numbers in Table 1, and so Fig. 1 has been prepared to show that the difference between the high- and low-probability performances is greatest during the last hour of the watch. Taking the first 30 min. as a whole, there was little difference between the two areas, and (at least for the CRT display) the targets in the low-probability area may actually have been easier to detect. If one thinks of such a decline in performance as a matter of relaxed vigilance (**2**), then one can say that *S*'s vigilance in the improbable areas tended to decline, whereas his vigilance in the probable areas remained relatively constant or even improved as the watch continued.

DISCUSSION

It would be desirable to have several consecutive days' testing on the same *S*s in order to see if the pattern established during the last hour of the first watch would persist into later watches or would have to be re-established each day. Even without such data, however, it seems obvious that a radar operator can be expected to look where the targets have been most probable. In some cases this influence of his subjective probabilities on his search behavior may be an advantage, but one can also imagine cases where it might be a disadvantage. For instance, if a radar operator is posted on the seacoast and told to look for planes out over the ocean, there may be many targets flying over land where he is not required to look and very few flying over water where he should be looking. Unless special precautions are taken, the operator will probably be less vigilant in the low-probability area that he is supposed to be monitoring.

It should be pointed out that the temporal decline in sensitivity for targets in the improbable areas involves something more than simply learning about the distribution of targets. There is no a priori reason why the observer should be less sensitive just because the targets are less frequent. It may be that as the watch proceeds, the operator gives up his effort to be eternally vigilant and relaxes in just those times and places that will hurt him the least—that is to say, where the targets are the least probable.

Two other experiments conducted with the films confirmed the general relations reported here: target detection was best where the targets were most probable, and this difference was far more obvious during the last hour of the watch than during the first 30 min.

The intermittent brightening and fading of the CRT display as the strobe line rotates over it is probably not the cause of the decline in performance after the first 15 min., since essentially the same decline is observed with the bright, steady, photographic display. The operator is not "hypnotized" by the strobe line. It is true that significantly more targets were detected on the photographic display, but the number of false reports was also much greater. This difference is attributed to the greater contrast on the film, which tended to make both the targets and the random clustering of the noise stand out from the background. It is concluded, therefore, that a photographic display,

despite its steadiness and brightness, does not prevent or even postpone the decline in detection that can be expected when the probability of a target is quite low.

In his theoretical discussion of vigilance phenomena, Deese (1) describes essentially the same experiment as reported here. His results, however, were negative. Deese speculates that his spatial distribution was not extreme enough to produce the effect, and, in the light of the present results, we are inclined to agree. If the difference in the probability of a target is made large enough to produce a genuine difference in the operator's expectations, his sensitivity reflects this difference.

Summary

When targets are much more probable in one area of a radar display than in another, there is a decline in sensitivity for the low-probability area and no decline for the high-probability area.

References

1. Deese, J. Some problems in the theory of vigilance. *Psychol. Rev.*, 1955, **62**, 359–368.
2. Mackworth, N. H. Researches on the measurement of human performance. *Med. Res. Council, Spec. Rep. No. 268*, London: His Majesty's Stationery Office, 1950.

(Received February 14, 1956)

Journal of Experimental Psychology
Vol. 53, No. 3, 1957

THE ROLE OF CLASS-DESCRIPTIVE CUES IN PAIRED-ASSOCIATES LEARNING

J. JEPSON WULFF[1] AND LAWRENCE M. STOLUROW[2]

University of Illinois[3]

The findings of several learning studies (e.g., **2, 3, 4, 5**) make it clear that stimulus processes are of basic importance to learning. The present experiment was designed to test the usefulness of one way to conceive of stimulus processes as they are related to cue utilization. The hypothesis is based upon a conception of the stimulus event in the S-R paradigm which differentiates between the stimulus object and the associative stimulus (s). Specifically, it is assumed that the associative stimulus, s, which is said to be associated with the "correct" response in paired-associates learning has three principal properties: (*a*) it is an "implicit" response to features of the stimulus object, the explicit or manipulated stimulus; (*b*) it has stimulus properties; and (*c*) it is a modifiable event rather than one that has a fixed relationship to the properties of the stimulus object. The associative stimulus is said to be modifiable because it is assumed that the specific "implicit" responses which function as the associative stimulus event in a given instance are in part a function of the learning conditions attending the formation of the association. Thus, if the stimulus objects in a paired-associates list permit the learner to make implicit responses to alternative features of the objects, one set of circumstances surrounding learning may favor one implicit response and another set of circumstances may favor a different set or even a complex of implicit responses as an associative stimulus. In this experiment the usefulness of this way of conceiving of the stimulus event was examined by comparing the level of learning resulting from two different training conditions. According to the above formulation, one of these training conditions would be expected to favor the utilization of combinations of cues, consisting of the cue which was both necessary and sufficient for correct response and an ancillary cue, as contrasted with the other condition favoring the utilization of an individual cue, which was necessary and sufficient for correct response. Furthermore, under the circumstances employed, the utilization of cue combinations of the type described would favor learning as compared with the utilization of an individual cue.

Problem.—Frequently, a list of paired associates may be subdivided descriptively into several subgroups by categorizing the stimulus terms on the basis of features which are the same within

[1] Now at Maintenance Laboratory, Air Force Personnel and Training Research Center, Lowry Air Force Base, Colorado.

[2] Now at Propulsion Research Unit, Maintenance Laboratory Field Unit Number 1, Air Force Personnel and Training Research Center, Chanute Air Force Base, Illinois.

[3] The material in this paper was presented at the twenty-sixth annual meeting of the Midwestern Psychological Association at Columbus, Ohio, in 1954. The assistance of Dr. R. Capobianco is gratefully acknowledged. This research was supported in part by the United States Air Force under Contract No. AF 33(038)-25726, monitored by the Air Force Personnel and Training Research Center. Permission is granted for reproduction, translation, publication, use and disposal in whole or in part by or for the United States Government.

stimulus-term subgroups but different between them. Features of stimulus objects that can be utilized in this way can be called *class-descriptive* features. For example, a list of paired associates with colored stimulus terms, each of which differs from all others in shape, may be divisible into subgroups of paired items on the basis of color. In such a list it is clear that color need not be attended to by the learner since shape alone will discriminate all of the stimulus terms. Nevertheless, with a long list of such paired associates one might expect that the color of each stimulus term would be attended to *in addition* to its shape, since to do so would have the effect of simplifying the task. Thus the long list would become a number of smaller tasks discriminable from each other on the basis of color. Under this condition, one would expect color to be especially helpful to the learner if a predifferentiated response class corresponded to each color or class-descriptive feature. For example, if the response terms for all of the yellow stimulus terms were numbers, and if no other stimulus term was paired with a number response, then yellow might be a useful class-descriptive feature of the stimulus terms to denote the "number" class of responses. However, in this case, since yellow could not be used alone as a cue for a specific response to any particular stimulus term, it would not be a necessary cue in learning to make correct responses to individual stimulus terms.

It is hypothesized that a class-descriptive feature of the stimulus terms, such as color in this example, may give rise to an implicit response that acts with the implicit response to each item-descriptive feature (a shape, for example)

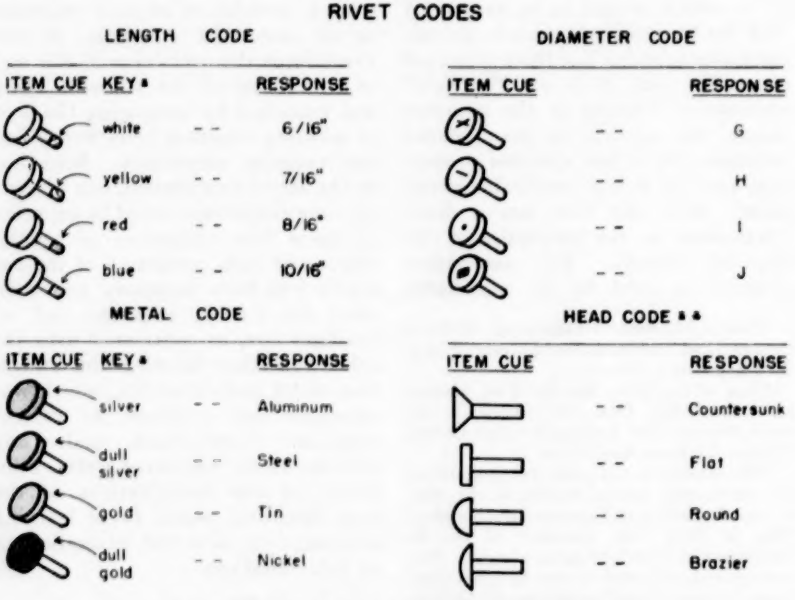

FIG. 1. Summary of the rivet code used in the experiment.

* While words are used in the present figure to label the differently marked features of these drawings, the actual colors signified by these words were used in the training and Test I materials of the experiment.

** Head type was not coded in the conventional sense of the word "code" since the shape of the side view of the manufactured head was employed as the cue; no abstract symbol was substituted for it as was done in the case of the other codes.

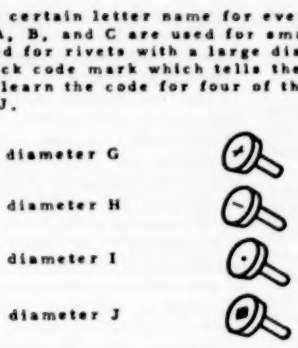
There is a certain letter name for every different rivet diameter. Letters like A, B, and C are used for small diameters; letters like X, Y, and Z are used for rivets with a large diameter. On the top of every rivet there is a black code mark which tells the diameter of the rivet. In this task you will learn the code for four of the most commonly used diameters: G, H, I, and J.
diameter G
diameter H
diameter I
diameter J

FIG. 2. A training card illustrating the *class*-organization method of presenting codes. This figure shows the presentation of the code for diameter.

as a compound stimulus event that is associated with the correct explicit response. If the implicit response to a class-descriptive feature of the stimulus terms has become associated, as a stimulus, with its response class, it can be expected to increase the probability of a correct response and decrease the probability of competing responses from other response classes through bond between the associative stimulus deriving from the class cue and the terms of the correct response class. It follows from this reasoning that learning conditions which favor the utilization of class cues (as well as item cues) should facilitate learning the total task as compared with those conditions which do not favor the use of class cues. The two methods of training employed in the present study were designed to permit a test of this hypothesis; one method favors the use of class cues, the other does not.

It should be noted that the term, cue, is used here to denote a feature of the stimulus object that gives rise to an associative-stimulus event.

METHOD

Learning Task

All *S*s used in the experiment were required to learn to interpret a specially devised code for rivets. As shown in Fig. 1, four different features of rivets were coded: length, diameter, type of metal, and type of head. Four different values of each of these four features were included in the task, making a total of sixteen paired associates in which the codes were treated as stimulus terms, and in which the values or names assigned to the codes were treated as response terms. Subsequent to experimental training by one of the two training methods used, *S*s were required to decode and to encode rivets in order to measure the amount of learning accomplished under the training methods used. One of these methods was designed to foster the use of class-descriptive cues, the other was not.

Training Materials

Inspection of Fig. 1 will reveal that the codes were designed in such a way that both pre-differentiated class-descriptive and item-descriptive cues are available for both stimulus and response terms. For example, it can be noted in Fig. 1 that "color on tip of rivet," is a class-descriptive cue for one subset of four stimulus terms, and that on the response side "length" responses or "16ths-of-an-inch" responses are appropriate for this class cue. Within this class of stimulus terms, the specific colors—white, yellow, red, and blue—serve to provide item-descriptive cues. Each color denotes a specific length response and is itself a necessary and sufficient cue for that response. The three other subsets of stimulus terms can be defined by the following class-descriptive cues: black mark on rivet head, metallic color of rivet head, and shape of side view of rivet head. Each of these three class cues denotes a different response class (i.e., diameter, metal, and head type, respectively), and within each of these subsets of stimulus terms there are four item-descriptive cues, each of which is uniquely related to a different response item. All item-

and class-descriptive cues were assumed to be minimally confusable to the *Ss*.

Experimental Conditions

In order to implement the experimental comparison, this code was taught by two different methods of presentation: a *class*-organized method and an *object*-organized method.

Class organization of materials.—This method of presentation was designed to afford conditions that would foster the use of class-descriptive cues. To this end the four paired associates for the length code (e.g., white—6/16 in., yellow—7/16 in., red—8/16 in., and blue—10/16 in.) were all presented together, and the same procedure was used for diameter, metal, and head type. Thus there were four 5 × 8-in. training cards, each with a different class of four paired associates. An example of a card constructed in this way is shown in Fig. 2.

Object organization of materials.—This method of presentation was designed to afford conditions that would not foster the use of class-descriptive cues. In this method each of four training cards displayed a different picture of a sample rivet. Every picture showed one cue from each of the four code classes. With this method, the information about the codes is organized into subtasks about a sample object, rather than about a rivet-feature class. The materials for *object*-organization training were prepared by selecting four sample rivets, each with four different code examples so that all sixteen code items were used. The coded features of each rivet picture were labeled in the manner indicated in Fig. 3. It can be seen that the balanced condition—four code subtasks, each of which contained four cue-response items—made it possible to equate the *class*- and *object*-organized training with respect to number of items presented at one time.

The first of these two arrangements, *class* organization, presumably favors the association of each class of responses with a separate class-descriptive cue. Thus, when *class* organization is used, the association of the class cue, "color on the tips," can readily be associated with all of the responses in the length class, 6/16, 7/16, 8/16, and 10/16 in. In contrast, with *object* organization of the training materials this type of association can take place only if the learner can recall (or anticipate) all the response terms for each class as each sample rivet is presented.

Training cards.—The training cards for each method of presentation were arranged into a pack of 10 cards. First, there were two introductory cards, then there were four cards explaining the four codes (as shown in Fig. 2 and 3) and finally, there were four review cards. The review cards did not include the verbal heading shown for the two cards in Fig. 2 and 3; otherwise, they were the same. The two introductory cards were exactly the same for both methods of training. The first contained instructions to the trainee for the use of the cards; the second explained what was meant by the length of a rivet, by the diameter of a rivet, and by the metal and head type of a rivet.

Test Materials

Two tests were used. One required *S* to interpret the codes applied to sample rivets (decoding); the other required them to tell how

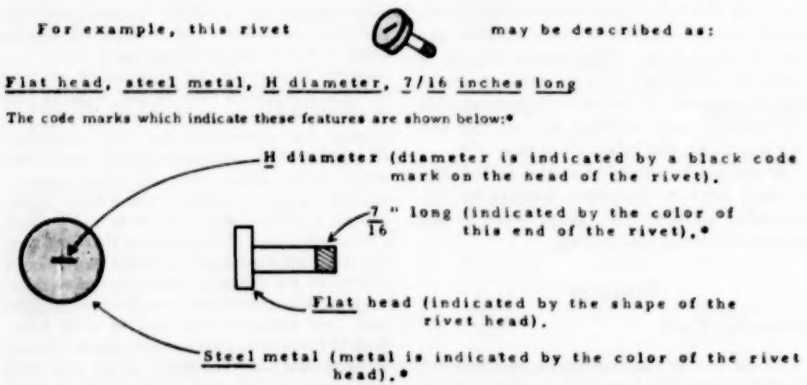
For example, this rivet
may be described as:
Flat head, steel metal, H diameter, 7/16 inches long
The code marks which indicate these features are shown below:*
H diameter (diameter is indicated by a black code mark on the head of the rivet).
7/16 " long (indicated by the color of this end of the rivet).*
Flat head (indicated by the shape of the rivet head).
Steel metal (metal is indicated by the color of the rivet head).*

FIG. 3. A training card illustrating the *object*-organization method of presenting codes.

* The training materials used in the experiment contained actual colors where shading is used in this figure, i.e., the head was dull silver and the tip was yellow.

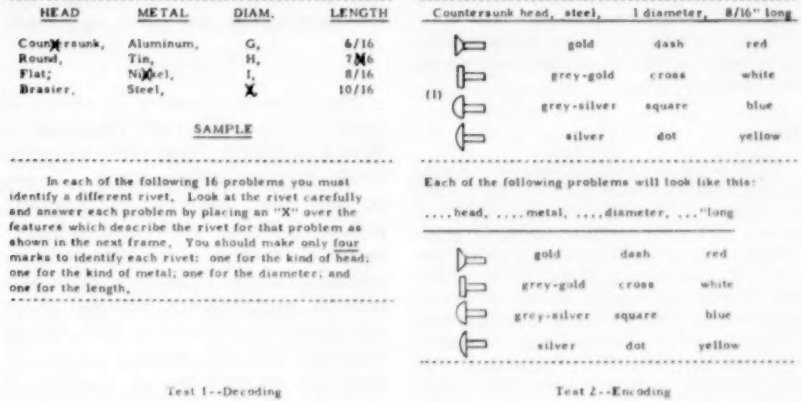
HEAD METAL DIAM. LENGTH
Countersunk, Aluminum, G, 6/16
Round, Tin, H, 7/16
Flat; Nickel, I, 8/16
Brazier, Steel, J, 10/16
SAMPLE
In each of the following 16 problems you must identify a different rivet. Look at the rivet carefully and answer each problem by placing an "X" over the features which describe the rivet for that problem as shown in the next frame. You should make only four marks to identify each rivet: one for the kind of head; one for the kind of metal; one for the diameter; and one for the length.
Test 1--Decoding
Countersunk head, steel, 1 diameter, 8/16" long
(1)
gold dash red
grey-gold cross white
grey-silver square blue
silver dot yellow
Each of the following problems will look like this:
...,head, ...,metal, ...,diameter, ..."long
gold dash red
grey-gold cross white
grey-silver square blue
silver dot yellow
Test 2--Encoding

FIG. 4. Instructions and a sample item for Test 1 and for Test 2, both of which were printed on a scroll. Only material between dashed lines was exposed at any one time. In Test 1, *S* decoded 16 actual rivets displayed in a tray; in Test 2, *S* indicated, also by crossing out one item in each answer column, how to encode 8 different rivets.

rivets with given characteristics would be coded (encoding).

Test 1—decoding rivets.—The first test required *S* to interpret 16 rivets coded in such a way that each cue was used four times but no pattern of three cues was represented more than once. To each of 16 rivets *S* made four responses.

As an experimental device, only two different actual lengths and two different actual diameters of rivets were used. Thus, the shorter-length rivets were coded to represent either 6/16 or 7/16 in., and rivets with the longer length were coded to represent 8/16 or 10/16 in. Similarly, smaller diameter rivets were coded to represent either I or J, etc. This was done to force *S*s to use the code rather than the perceived length or diameter of the rivet to determine their responses in describing the rivets. To eliminate weight cues all of the rivets were made of aluminum.

Each of the 16 test rivets was placed in one of the 16 numbered cells of a tray. Answer scrolls which fit into a reading-rate recorder—a machine which allowed *S* to expose in a window one answer panel at a time—were prepared with 16 numbered blocks or panels. Each panel contained all 16 of the possible response items grouped according to the four subtasks, i.e., by head, diameter, metal, and length. The *S* was instructed to pick up each rivet in turn and to cross out the four correct responses in the answer panel. A sample answer panel is shown in Fig. 4.

Test 2—encoding.—For the second test, eight more answer panels were prepared on the same scroll. At the top of each of these panels there was a description of a rivet in terms of the response items (for example: "Flat head, aluminum, G diameter 6/16" long,"), and below the rivet description all 16 code-cue stimuli were listed. They were arranged in four arrays as in Test 1 (see Fig. 4), and *S* crossed out the four appropriate code descriptions for each rivet. The *S*s made 32 responses in this test, four responses to each of eight rivet descriptions. Each of the 16 response names was used twice in this test, and no two rivet descriptions had as many as three items in common. The printed instructions for Test 2 (encoding) are presented in Fig. 4.

The *S*s paced themselves in taking both of the tests; there was no time limit.

Procedure

The order of events for each *S* was as follows: (*a*) read each of the 10 training cards once; (*b*) take Test 1; (*c*) take Test 2.

Before training began, *S*s were instructed to read the cards one at a time at their normal rate and *not* turn back to cards they had already finished. The total time spent in reading the training cards was measured by an ink pen-recording mechanism built into the reading-rate recorder. The machine's recording system

moved a pen horizontally across a scroll at a constant rate while *S* read a card.

Testing was begun as soon as the training cards had been finished, and no break was allowed between tests.

Design

Each of the two training methods was used to train a different group of 20 *Ss*. Half of the *Ss* trained by each method were Air Force trainees with high mechanical-aptitude scores ("High ability group") and half were trainees with low mechanical-aptitude ("Low ability group") scores as measured by the mechanical-aptitude index of the Airman Classification Test Battery (1). Two *Ss* were trained at a time. With two treatments this meant that four *Ss* were necessary to complete one block (replication) in order to equate the frequency within a block with which the two reading machines appeared. All *Ss* within an experimental design block had the same mechanical-aptitude index; two of the four were assigned to each treatment in an order determined by reference to a table of random numbers. Since *Ss* were used as they were available, this procedure resulted in an order of blocks which was not sequentially biased with respect to ability level. That is, all high-ability *Ss* were not run either first or last, but rather high- and low-ability blocks were intermixed. The design was therefore a two-by-two factorial (Training Method × Mechanical Aptitude Index) with 10 *Ss* in each of the four cells.

Results

The principal results in terms of mean number of correct responses in Test 1 and Test 2 are summarized in Table 1 by training method. It can be seen that *Ss* given *class*-organization training performed reliably better than those given *object*-organization training. The over-all difference in performance between these two groups on Test 1 is 9.65 correct responses which is reliable at the .01 level, as assessed by an analysis of variance (Table 2). Similarly, the difference between these two training conditions on Test 2 of 5.35 is reliable ($P < .05$).

As indicated above, high and low mechanical-aptitude groups were included in the design. This dichotomy was included to increase the precision of the experimental design rather than because of any primary interest in performance as related to aptitude

TABLE 1

Mean Number of Correct Responses for Test 1 and Test 2 By Experimental Conditions

Test	Variable	Group	Type of Organization of Materials		Reliability of Differences
			Class Organization ($N = 20$)	Object Organization ($N = 20$)	
Test 1 (64 items)	Aptitude	Over-all	48.0	38.35	$P < .01$
		High	54.9	46.2	$P = .05$
		Low	41.1	30.5	$.05 > P > .02$
	Subtask (16 items)	Head	13.95	13.70	*
		Metal	9.30	7.20	*
		Diameter	12.20	9.25	*
		Length	12.55	8.20	*
Test 2 (32 items)	Aptitude	Over-all	24.65	19.80	$P < .05$
		High	27.3	24.8	$P < .05$
		Low	22.0	14.8	

* P values are not appropriate for these comparisons since the differences are not independent of the values presented above.

TABLE 2

SUMMARY OF ANALYSES OF VARIANCE

Source	*df*	Test 1		Test 2	
		MS	*F*	MS	*F*
Trng. Method (M)	1	931.22	7.68**	235.22	6.40*
Aptitude (A)	1	2,175.62	17.96**	585.22	15.94**
A × M	1	9.02		55.22	1.50
Within	36	121.10		36.70	

* $P < .03$.
** $P < .01$.

level. Nevertheless, the aptitude groups were compared separately as shown in Table 1, and the performance difference between training conditions is found to be reliable for each of the aptitude groups

The mean number of correct responses per subtask is also shown in Table 1 by training method for Test 1. It can be seen from the tabled means that the superior performance of students trained by the *class*-organization method is consistent for all subtasks, indicating that the over-all comparison does not reflect merely a large superiority with respect to one or two subtasks.

The difference in mean speed of reading the training materials is small and unreliable for the two training groups (class training—320.49 sec.; object training—357.66 sec.). This finding makes the interpretation of the test results clear. Since the two groups spent about the same amount of time studying the same content differently organized, it is apparent that organization of the training materials rather than reading time is related to the obtained differences in test performance.

DISCUSSION

The results clearly demonstrate that with these experimental materials *class*-organized training materials are more effective than materials organized according to the object method. This outcome is consistent with the expectation that class-organized training conditions would favor the development of associative-stimulus events involving implicit responses to class cues which would in turn foster learning the total task. Thus one may interpret these results in terms of the associative-stimulus concept stated above in which it is assumed that the stimulus event, s, is modifiable.

If one chose to interpret the present findings in terms of class and item cues, one must recognize that the role of the class-descriptive cue can be formulated in alternative ways which are compatible with these data. As it was stated, the hypothesis tested in this experiment does not provide a basis for differentiating between the two alternative paradigms that could be postulated to account for the learning process involved.

One of these paradigms would assume a dual habit structure. For example, an associative stimulus (s^c) based upon a class cue may become associated with all of the response terms (R^cs) for that class as though the list of response terms were a single response. An associative paradigm for this habit can be represented for Class 1 as indicated below.

$$s^{c-1} - \begin{cases} R_{1-1}^{c-1} \\ R_{1-2}^{c-1} \\ R_{1-3}^{c-1} \\ R_{1-4}^{c-1} \end{cases}$$ (First habit in a hypothetical dual habit paradigm)

The second habit would be a simple association between the correct response term and the implicit response to the item cue (e.g., for item one, s_{i-1}-R_{i-1}; for item two, s_{i-2}-R_{i-2}, etc.). In the convention adopted to represent the habit elements, the superscript $^{c-1}$ denotes Class 1 and the subscript $_{i-1}$ denotes Item 1. In this experiment, Class 1 might be "color on rivet tip" and the associated responses for the first habit "lengths 6/16, 7/16, 8/16, and 10/16 in." For the second type of habit an s_{i-1} event might be "white" and its response "6/16 in." As indicated with this assumed type of associational structure, item cues would be responded to as a separate step in performance

following the response to the class-descriptive cue.

The role of class-descriptive cues in the dual habit paradigm may be contrasted with a process in which a single unique implicit response to each combination of class and item cues is acquired. This state of affairs can be represented by the following associative paradigm—s_{i-1}^{e-1}-R_{i-1}^{e-1}. From the point of view of this second paradigm, the stimulus event is a single unique compound of class- and item-descriptive cues. The stimulus event for the same example as used above would be "white color on rivet tip" (s_{i-1}^{e-1}) and the associated response "length, 6/16 in." (R_{i-1}^{e-1}).

If the latter condition existed, one would expect positive transfer of stimulus learning in an A-B, A-K transfer situation involving class-descriptive cues. Positive transfer of this type has been found by Wulff (**5**) and by Kurtz (**2**). Wulff's experiments, however, did not involve response subclasses. In Kurtz's experiments, the item-descriptive cues were obscure, which condition encouraged the learner to attend first to class cues; and one might hypothesize that with obscure item cues, class training would be facilitating for a different reason from that assumed in the present study. Thus in Kurtz's study the class conditions would force the learner to attend to item cues early in learning. This explanation does not apply to the present training situation. The available data do not permit a decision regarding which of the two types of habit structure the class training induced.

Summary

It was postulated that the stimulus unit represented by S in the associative paradigm can be usefully described as: (*a*) an "implicit" response to features of the stimulus object; (*b*) an event with stimulus properties; (*c*) a modifiable event rather than one that is fixed by the stimulus object. It was hypothesized that when the stimulus objects in a list of objects have class-descriptive features, these features may give rise to implicit responses which become a part of the associative-stimulus event. In this case the associative-stimulus event is said to be compounded of implicit responses to class-descriptive cues and item-descriptive cues.

The utility of this conception of the stimulus unit was tested by comparing the rate of paired-associates learning for a given task under two different conditions of training. The paired-associates task was one for which learning was expected to be facilitated if combinations of implicit responses to class and item cues functioned as the stimulus events. Two training conditions were selected so that one would favor the formation of such compound stimulus events, the other would not. The training condition which favored the use of combinations of cues resulted in a reliably higher level of performance with a given amount of training.

References

1. Daily, J. T. Development of the airman classification test battery. San Antonio, Tex.: Human Resources Research Center, Lackland Air Force Base, November 1948 (*Res. Bull.* 48-4).
2. Kurtz, K. H. Discrimination of complex stimuli: the relationship of training and test stimuli in transfer of discrimination. *J. exp. Psychol.*, 1955, **50,** 283–292.
3. Osgood, C. E. *Method and theory in experimental psychology.* New York: Oxford Univer. Press, 1953.
4. Spence, K. W. The nature of discrimination learning in animals. *Psychol. Rev.*, 1936, **43,** 427–449.
5. Wulff, J. J. The stimulus unit in associative learning. Unpublished Ph.D. dissertation, Yale Univer., 1954.

(Received February 10, 1956)

Journal of Experimental Psychology
Vol. 53, No. 3, 1957

A DEDUCTION FROM AN EXCITATION–INHIBITION ACCOUNT OF RETROACTIVE INHIBITION [1]

ERNST Z. ROTHKOPF [2]

Maintenance Laboratory, Air Force Personnel and Training Research Center, Lowry Air Force Base, Colorado

Several theoretical formulations (e.g., **2, 6, 7, 8**) propose an excitation-inhibition schema for learning. This theoretical schema is characterized by three major mechanisms. These are: (*a*) increases in excitatory strength contingent upon reward or confirmation of a stimulus-response sequence; (*b*) inhibition of unrewarded or unconfirmed responses; and (*c*) generalization of excitatory and inhibitory strength along one or more gradients of stimulus similarity. In addition, it is useful to assume that excitation or inhibition will take place regardless of whether the relevant response is *overt* or *covert*.

The excitation-inhibition paradigm appears of value in a theoretical account of retroactive inhibition. According to the schema described above, the retroactive inhibition effect, i.e., the decrement in S_1-R_1 (the original list) following practice with S_1-R_2 (the interpolated list) is explainable in the following terms. First, the growth in the excitatory strength of S_1-R_2 results in increased competition for S_1-R_1 and thus decreases the probability that S_1 will elicit R_1. Second, the observable strength of S_1-R_1 diminishes because of the inhibition of R_1 due to its unreinforced elicitation during practice with the interpolated list. This application of an excitation-inhibition paradigm to the retroactive inhibition effect is, of course, essentially congruent with the two-factor account of retroactive inhibition proposed by Melton and associates (**4**).

The present theoretical schema generalizes beyond the so-called classical retroactive inhibition situation—A-B, A-K, A-B—to the case in which the stimulus terms of the interpolated paired-associate list are not identical with the stimulus terms of the original list, i.e., A-B, S-K, A-B. In the latter situation it is assumed that under certain conditions (between-list stimulus similarity) the stimuli of the original list (S_1) generalize to stimulus terms of the interpolated list (S_2). This generalization results in the *overt* or *covert* occurrence of first-list responses (R_1) during practice with the second list. Because the occurrence of S_2-R_1 is not reinforced under these circumstances, this response tendency is inhibited. The inhibition of R_1 then generalizes from S_2 to S_1, as does the excitation of R_2. The consequent rise in the inhibition of S_1-R_1 and in the excitatory strength of S_1-R_2 diminish the observable strength of S_1-R_1 and so produce the retroactive inhibition effect.

The theoretical account elaborated above depends to an important degree upon the supposed inhibition of S_2-R_1 during practice with the interpolated or second list. A direct implication of this notion is that *after* two paired-associate lists with similar stimuli are practiced in succession and when other factors are controlled, *the tend-*

[1] The experimentation reported here was part of exploratory research carried on in support of Project 7714 of the United States Air Force Personnel and Training Research Program.

[2] The author is indebted to A/1C Richard H. Perkins and A/2C John M. Mayhall III, who performed the actual experimental work with great skill.

ency to erroneously give second-list responses to first-list stimuli (S_1-R_2) *should exceed the tendency to make errors involving first-list responses given to stimuli of the second list* (S_2-R_1). In concordance with above theoretical schema this is to be expected because even though both types of error tendencies rise through generalization processes during practice with the appropriate list, only these between-list errors which involve *first-list responses* given to *second-list stimuli* (S_2-R_1) are inhibited. Data appropriate to a test of this deduction can be obtained by comparing the frequency of between-list errors for *Ss* who practice two paired-associate lists in opposite order. The *Ss* who train in the sequence S_A-R_A, S_B-R_B should show fewer S_B-R_A errors on a test administered subsequent to practice with the two lists than *Ss* who train in the sequence S_B-R_B, S_A-R_A. This is to be expected because *Ss* who practice S_A-R_A, S_B-R_B, start practice with the second list when S_A-R_A is of appreciable strength. Therefore, because of generalization processes, many overt or covert S_B-R_A errors occur during second-list practice and S_B-R_A is therefore inhibited. In the case of *Ss* who practice S_B-R_B *before* S_A-R_A is established, little if any inhibition of S_B-R_A is to be expected because R_A is unlikely during practice with S_B-R_B. (It follows, of course, that the S_B-R_B, S_A-R_A sequence group should show a smaller number of S_A-R_B errors.) The present study was designed to test this deduction for between-list errors under the conditions described above.

Method

Procedure.—Prior to the experiment proper, all *Ss* were given eight practice trials with a paired-associate list consisting of two-syllable Russian nouns as stimulus terms and short English nouns as response terms. The list was composed of six paired-associates of this type. This pretraining task was intended to accustom *Ss* to the experimental environment and to familiarize them with a paired-associate procedure akin to that used in the main experimental task.

The experimental task consisted of practice in writing a different letter of the alphabet in response to each of 12 signals of the International Morse Code. These signals were: · – – (W), – – · (G), · – · (R), · – – – – (1), · · – – – (2), · · · – – (3), · – – · (P), · – · (F), · – · · (L), · · · (S), · · · · (H), · · · · · (5). The appropriate alphabetic equivalents were used as response terms except that to prevent any special effects due to "numbers" as a response class, the letter responses D, N, V, and K were paired with the number signals 1, 2, 3, and 5, respectively. The signals were split into two lists of six paired-associates each. List A consisted of: W, 1, 3, P, S, 5. List B consisted of: G, R, 2, F, L, H.

These lists were composed so that List A contained stimuli which were similar to stimulus terms of List B. Similarity among the stimuli within each list was also relatively high. Similarity between two stimuli for this purpose was the percentage of the total presentations in a psychophysical situation with which these two stimuli were judged by naive *Ss* as identical (see **5**). This measure has been shown to be positively related to the frequency of substitution errors in paired-associate learning (**5**). Average within-list stimulus similarity using this measure was 17.7%. Average between-list similarity was 24.3%. The median percentage of "same" judgments for objectively different Morse stimuli is 13%.

Practice was conducted by means of the code-voice method. This technique, an adaptation of the paired-associate method, is described in detail elsewhere (**3**). In brief, the version used here consisted of the presentation of a signal, followed by a 3-sec. silent period during which *S* could write his answer. This period was in turn followed by the signal's spoken-letter equivalent. The next Morse stimulus followed 1.6 sec. later.

Each *S* received 60 practice trials with each list. A practice trial consisted of a single, random-order presentation of each of the six paired associates in a list. A six-signal test run was interposed after every fourth practice trial. Test runs were composed in the same manner as practice trials except that the signals' spoken-letter equivalents were omitted.

Immediately after the last practice trial with the second list, a 48-item test was administered to all *Ss*. This test consisted of four presentations of *each of the 12 stimuli* in random order.

Twenty-one hours after the administration of this test, a second, similarly composed test was given to all *Ss*.

Materials and apparatus.—The signals of the International Morse Code used in this experiment were produced using a Boehme Automatic Keyer at a tone speed of 20 words per minute. This tone speed means that each signal was sent fast enough to permit 20 transmissions of the word "Codez" per minute, or that the temporal duration of each dot was .05 sec., that each dash was .15 sec., and that a silent period between elements was .05 sec. A male voice announced the signal's letter equivalent using the International Phonetic Alphabet. The pretraining materials, as well as the materials for experimental practice and tests, were recorded on magnetic tape.

The experiment was carried out in a windowless, air-conditioned classroom. Each of the 28 seats in the room was equipped with a pair of earphones. The output of any one of several tape recorders could be channeled to each headset. This permitted the concurrent administration of different experimental treatments to *Ss* sitting in the same room.

Subjects.—The experimental population consisted of 55 airmen awaiting entrance into an Air Force Radio Operator School. None of the *Ss* indicated that they had had previous Morse code experience.

Design.—The *Ss* of this experiment were used in two groups of 27 and 28 men each. Pretraining and testing was the same for all *Ss*. Approximately half of the *Ss* of each group practiced first with List A and then with List B. The remaining *Ss* practiced the lists in opposite (B, A) order.

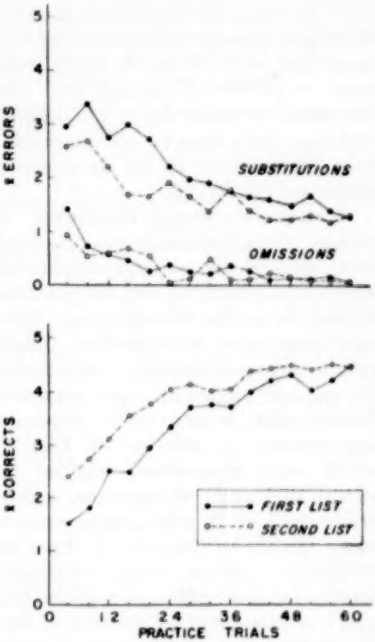

FIG. 1. The incidence of correct responses, omissions, and substitution errors as a function of practice. The solid circles represent the mean of data from both Lists A and B when they were practiced first in the training sequence. The hollow circles are like data for the second position in the training order.

RESULTS

Results will be summarized in terms of two sets of data: (*a*) acquisition scores based upon the 6-stimuli tests which were periodically administered throughout practice with either list; and (*b*) postpractice test scores based upon a 48-item randomized presentation of *all 12 stimuli* 0 and 21 hr. after practice with the second list.

Acquisition.—The 6-item tests which were administered after every four practice trials were scored for correct responses, omissions, and substitution errors of any sort. Changes in these three response categories are shown graphically in Fig. 1. In this figure the acquisition data from the list which was practiced first are contrasted with like data from the list which was practiced second. Approximately half[3] of the *Ss* which contributed to each data point in Fig. 1 practiced with List A, the remainder with List B. Figure 1 indicates that *Ss* made fewer substitution errors and more correct responses during early practice with

[3] A slight inequality resulted from the fact that one *S* failed to arrive at the appointed hour. As a consequence, 27 *Ss* first practiced List A and then List B, while 28 *Ss* practiced in the B, A order.

their second paired-associate list than they did during early practice with their first list. After 60 trials, however, *Ss* performed equally well with the items of either list. Throughout training, there were no differences due to practice order so far as omission errors are concerned.

Postpractice tests.—Responses to stimuli of the list which was practiced first, and to the stimuli of the list which was practiced last, were scored separately on the 48-item tests. Correct responses, between-list substitution errors, within-list substitution errors, and omissions were recorded. Between-list errors were responses appropriate to stimuli of List A which were erroneously elicited by stimuli of List B, or responses appropriate to List B which were erroneously given to stimuli of List A. Within-list errors were responses, correct for a stimulus in a given list, which were elicited by another stimulus of that list. Failure to respond and illegible responses were scored as omission errors. Ten of the responses which were scored (i.e., .19%) could not be fitted into the categories enumerated above. These were counted as omission errors.

For brevity, the notations described below will be used to refer to the various categories of responses throughout the remaining sections of this paper. Between-list errors made by *S* to stimuli of the list which he practiced first, regardless of whether List A or List B materials were involved, will be designated S_1-R_2 since they involve responses appropriate to the second list. The subscripts indicate in each case the position in the practice order of the paired-associate list to which the stimulus or the response is proper. Analogously, between-list errors which consist of first-list responses to stimulus terms from a list which was second in a *S*'s practice order will be designated S_2-R_1. Correct responses to stimuli of the first and second list will be designated S_1-R_1 and S_2-R_2, respectively. Within-list errors elicited by stimuli of the first-practiced list will be written as S_1-R_{1E} and those elicited by stimulus terms of the second list as S_2-R_{2E}. Omission errors will be termed in accordance with the usage described above as S_1-R_O and S_2-R_O, respectively.

The mean occurrence of each of the various responses per stimulus for the 48-item test which was administered immediately after the last practice trial is shown in the 0-hr. row of Table 1. Since each stimulus was presented four times during each 48-item test, the mean frequencies in the several response categories sum to *four* in the rows. As predicted, S_2-R_1 errors were less frequent than S_1-R_2 errors. Of the within-list errors S_2-R_{2E} responses were somewhat more frequent than S_1-R_{1E} responses. The table also indicates that the S_2-R_2 frequency was greater than S_1-R_1.[4] All three of these differences were significant beyond the .01 level as tested by Wilcoxon's paired-replicate procedure[5] (**11**, p. 5). The paired replicates for this purpose were the number of times a given stimulus elicited a particular type of response when the stimulus appeared first in the training order and the number of times that the same stimulus elicited the same type of response when the

[4] On the six-item test which followed the 60th practice trial, the stimuli from the first list elicited approximately the same number of correct responses and substitution errors as the stimuli of the second list. The apparent discrepancy between these data and the 48-item test under discussion resulted from the way the practice situation was structured. The instructions which preceded each 6-item test stated that *Ss* were being tested on the materials which they had just practiced. This was a sufficient stimulus to prevent any *overt* between-list errors from occurring during these tests. On the 48-item test the instructions indicated that *Ss* would now be tested on all 12 paired associates. Consequently all twelve responses were "legitimate" and between-list errors occurred frequently. Only the 48-item test data reflects the competition of all 12 response tendencies.

[5] This statistical procedure was used because the distributions of all four categories of responses were skewed.

TABLE 1

COMPARISON OF MEAN FREQUENCIES OF THE FOUR RESPONSE CATEGORIES ON TESTS ADMINISTERED 0 AND 21 HR. AFTER PRACTICE FOR STIMULI OF THE LIST WHICH APPEARED FIRST IN THE PRACTICE ORDER AND OF THE LISTS WHICH APPEARED SECOND

(The measures shown are the average response frequencies per *S* for the four presentations of each stimulus term during a 48-item test.)

Interval Between Practice and Test	Between-list Errors			Within-list Errors		
	List		Diff. (2nd − 1st)	List		Diff. (2nd − 1st)
	1st (S_1-R_2)	2nd (S_2-R_1)		1st (S_1-R_{1E})	2nd (S_2-R_{1E})	
0-hr.	1.85	.74	− 1.11**	.45	.83	.38**
21-hr.	1.40	1.07	− .33	.58	.61	.03
Diff. (21-hr.−0-hr.)	− .45**	.33**		.13	− .23*	

Interval Between Practice and Test	Omission Errors			Correct Responses		
	List		Diff. (2nd − 1st)	List		Diff. (2nd − 1st)
	1st (S_1-R_O)	2nd (S_2-R_O)		1st (S_1-R_1)	2nd (S_2-R_2)	
0-hr.	.83	.80	− .03	.87	1.62	− .75**
21-hr.	.84	.80	− .04	1.18	1.52	− .34*
Diff. (21-hr.−0-hr.)	.01	0		.31**	− .10	

* P = .05, using Wilcoxon's paired-replicates test.
** P = .01, using Wilcoxon's paired-replicates test.

list in which the stimulus appeared was last in the training order. For example, the mean number of between-list errors made to the stimulus ····· (5) for *S*s who trained with List A (W, 1, 3, P, S, 5) first, were compared with the mean number of between-list errors made to the same stimulus ····· (5) by *S*s who trained with List B (G, R, 2, F, L, H) first, and so on for each of the 12 stimulus terms used in this experiment. When this statistical test was applied to omission errors, no statistically significant differences were obtained.

Comparable data for the 48-item test, which was administered 21 hr. after the end of formal practice, is also shown in Table 1. Although S_2-R_2 frequency was still significantly greater than S_1-R_1, this difference is only about half as large as in the test administered immediately after practice. This advantage as far as correct responses to the second-position list were concerned was of about the same order of magnitude as the difference between the S_2-R_1 and the S_1-R_2 frequencies. S_2-R_1 was still less frequent than S_1-R_2. However, this difference was much smaller than it was on the test administered immediately after practice and did not reach statistical significance. Differences in the number of within-list and omission errors due to order of training were small and not statistically reliable.

In order to permit a more detailed evaluation of the changes which took place during the 21-hr. postpractice interval, the mean frequencies of each

of the various response categories before and after the 21-hr. rest were compared. Wilcoxon's paired-replicate procedure was again used. The paired replicates were the incidence of a given category of response (e.g., S_1-R_1, S_2-R_2, etc.) before and after the rest period for *each stimulus*. The results are shown in the Diff. (21-hr. − 0-hr.) rows of Table 1. Four of the eight possible comparisons were significant beyond the .05 level. These were S_1-R_1 and S_2-R_1 which increased in average frequency during the 21-hr. rest interval, and S_1-R_2 and S-$_2R_{2E}$ which decreased during that time period.

Discussion

The major finding of this experiment was that *Ss* who were tested with the stimulus terms of two recently practiced paired-associate lists made fewer between-list errors to stimulus terms of the last-practiced list than they did to stimulus terms of the first-practiced list (S_2-$R_1 < S_1$-R_2). Since both lists were given equal practice and were in other respects comparable, the results of this study support the notion that first-list response tendencies[6] are inhibited during practice with the second list. Such inhibition is assumed to result from the nonreinforced covert elicitations of first-list responses during practice with the second list. This notion was developed in the introductory portions of this paper.

The relative frequency of correct responses as well as within-list error responses are also relevant to the major hypothesis under test. The findings of the present experiment in this respect support the hypothesis that first-list response tendencies are inhibited since S_2-R_2 were found to be more frequent than S_1-R_1, and S_2-R_{2E} are more frequent than S_1-R_{1E}. No method is known at present for determining to what extent the fact that second-list stimuli elicit correct responses more often than first-list stimuli is due to (*a*) the reduced competition of S_2-R_1 tendencies, and (*b*) the generalization of inhibition of R_1 from S_2 to S_1 and the consequent weakening of the S_1-R_1 tendency. The relative contribution of analogous factors to the greater number of within-list errors among the stimuli of the second list is similarly indeterminate.

Twenty-one hours after practice the difference between the first and the second position in the practice order as far as the frequencies of the various response categories are concerned has diminished to a considerable degree. This is consistent with the results of two studies by Underwood (**9, 10**). This investigator found that the second of two consecutively practiced paired-associate lists is retained better than the first after 5 hr., but that both lists were equally well retained 48 hr. after practice (**9**). Underwood also presents evidence (**10**) which suggests that first-list responses undergo something like "spontaneous recovery" during a 48-hr. interval between learning and recall. Similarly Briggs (**1**) reports that the frequency of first-list responses increases during the early portions of a 72-hr. retention interval. If it is assumed that the inhibition of first-list responses dissipates in time, the Underwood and Briggs material as well as the present data are respected.

Specifically in the case of the present study, the dissipation of the inhibition of R_1 will result in increments in the strength of S_1-R_1, S_1-R_{1E}, and S_2-R_1. Increases in the strength of these three response tendencies would result in smaller differences between S_2-R_1 and S_1-R_2 than were observed immediately after practice because increments in

[6] In discussing the findings of this experiment, a distinction is made between *response frequency* and *response tendency* or *strength*. The first term—response frequency—refers to the number of times a stimulus was observed to be followed by a particular response or response category. This term is analogous to Hull's effective reaction potential, $_sE_R$ (**2**). Response tendency or strength, on the other hand, is an unobservable which can be understood as a stimulus' potential for eliciting a particular response or class of responses. Response tendency is analogous to Hull's effective habit strength, $_s\hat{H}_R$ (**2**).

S_2-R_1 would increase the observable strength of S_2-R_1, and also because increased competition of S_1-R_1 and S_1-R_{1E} should lower the observable strength of S_1-R_2.

Increases in the strength of S_2-R_1 should also lower the observable frequency of S_2-R_2 and S_2-R_{2E}. Both of these response frequencies change in the predicted direction during the 21-hr. rest interval but only the decrease in S_2-R_{2E} is statistically significant.

The data also indicated a statistically reliable increase in the observable incidence of S_1-R_1 during the 21-hr. postpractice interval. Although the S_1-R_1 tendency is expected to recover during the rest period, a tendency which competes with it, S_1-R_{1E}, is also on the increase. However, it is likely that the competition of S_1-R_{1E} is slight after the amount of practice given. Nevertheless, the increase in the frequency of S_1-R_1 could not be firmly based on the present theoretical schema. Quantitative assumptions about the gradients of excitation and inhibition which underlie the effects reported in the present study would increase the power of an excitation-inhibition schema for paired-associate learning to a considerable degree.

Summary

Fifty-five airmen practiced two paired-associate lists of six items each. The stimulus terms for these lists consisted of signals of the International Morse Code and the response terms of written letters of the alphabet. The experimental population was divided into two approximately equal groups, each of which practiced the two paired-associate lists in opposite order from each other.

When *Ss* were tested with all 12 paired associates immediately after practice with the second list, the performance data supports the following conclusions: (*a*) Between-list errors were more frequently made to stimulus terms of the first-practiced list than to stimulus terms of the list which was second in the practice order; (*b*) correct responses and within-list errors were more frequently made to the stimulus terms of the paired-associate list which was practiced last.

When *Ss* were tested 21 hr. after practice with the second list, stimuli of the list which was second in the practice order were responded to correctly more often than the stimuli of the first list. However, this effect had diminished to a considerable degree during the 21-hr. interval. As far as other response tendencies were concerned, no reliable differences were obtained between the stimuli of the first and the second list on the 21-hr. postpractice test.

These findings were interpreted as supporting the notion that retroactive inhibition depends to an important degree on the inhibition of overt or covert original list responses during practice with the interpolated list. The 21-hr. recall data suggest that these inhibitory effects decay in time. The general results of this experiment confirm the value of an excitation-inhibition account of paired-associate learning.

References

1. Briggs, G. E. Acquisition, extinction, and recovery functions in retroactive inhibition. *J. exp. Psychol.*, 1954, **47**, 285–293.
2. Hull, C. L. *Principles of behavior*. New York: D. Appleton-Century, 1943.
3. Keller, F. S. Studies in International Morse Code. 1. A new method of teaching code reception. *J. appl. Psychol.*, 1943, **27**, 407–415.
4. Melton, A. W., & Von Lackum, W. J. Retroactive inhibition in retention: evidence for a two-factor theory of retroactive inhibition. *Amer. J. Psychol.*, 1941, **54**, 157–173.
5. Rothkopf, E. Z. A measure of stimulus similarity and errors in some paired-associate learning tasks. *J. exp. Psychol.*, 1957, **53**, 94–101.
6. Saltz, E. A single theory for reminiscence, act regression, and other phenomena. *Psychol. Rev.*, 1953, **60**, 159–171.
7. Spence, K. W. Analysis of formation of visual discrimination habits in chimpanzee. *J. comp. Psychol.*, 1937, **23**, 77–100.
8. Spence, K. W. The differential response in animals to stimuli varying within a single dimension. *Psychol. Rev.*, 1937, **44**, 430–444.
9. Underwood, B. J. Retroactive and proactive inhibition after five and forty-eight hours. *J. exp. Psychol.*, 1948, **38**, 29–38.
10. Underwood, B. J. 'Spontaneous recovery' of verbal associations. *J. exp. Psychol.*, 1948, **38**, 429–439.
11. Wilcoxon, F. *Some rapid approximate statistical procedures*. New York: American Cyanamid Co., 1949.

(Received March 5, 1956)

Journal of Experimental Psychology
Vol. 53, No. 3, 1957

PERFORMANCE IN EYELID CONDITIONING FOLLOWING INTERPOLATED PRESENTATIONS OF THE UCS[1]

K. P. GOODRICH, L. E. ROSS, AND A. R. WAGNER

State University of Iowa

A recent study by Kimble, Mann, and Dufort (**2**) involving eyelid conditioning reports that repeated administration of the UCS alone after a series of paired-conditioning trials is functionally equivalent to an equal number of paired trials as revealed by subsequent conditioned responding. Specifically, an experimental group of *Ss* was given 20 CS-UCS pairings, then 20 trials with the air puff alone, and finally 20 more paired-conditioning trials. It was found that these experimental *Ss* gave a comparable or slightly higher percentage of CR's in Trials 41–60 as compared with *Ss* in the control group that received both CS and UCS throughout the 60 trials.

An interpretation of these results was presented which required two assumptions: (*a*) that the associative factor (habit), although a function of CS-UCS pairings, is acquired rapidly, reaching a maximum well within 20 trials, possibly within 3 to 5 trials, and (*b*) that a motivational factor (drive) continues to increase gradually in strength with repeated presentations of the UCS. Accepting these assumptions and an appropriate combination function, it is possible to predict a gradually rising performance curve of conditioning. The traditional Hullian view holds that the associative factor (*H*) reaches its maximum only after a considerable number of paired trials, and that drive is essentially constant in most conditioning situations. Thus, according to this interpretation, it is increasing habit strength, not increasing drive strength, that is reflected in the rising performance curve. The Kimble-Mann-Dufort design offers a test of the alternative predictions following from these two formulations.

Although the results discussed above do follow from the authors' assumptions, theoretical conceptions involving the notion of gradual habit growth have had considerable success in mediating deductions of eyelid phenomena in this laboratory and elsewhere (**4**). As shall be discussed later, the experimental procedures used in the Kimble, Mann, and Dufort study were different in a number of critical respects from those employed in this laboratory. Together with certain evidence which indicates that the motivational properties of the UCS may actually diminish with repeated presentations (**3, 6**), these considerations lead to skepticism regarding the generality of the findings of these investigators. The present study makes use of their experimental design, but employs the procedures and recording techniques used in this laboratory.

METHOD

Subjects.—Sixty women and 15 men from a summer course in introductory psychology served as *Ss*. Eight women who gave CRs to test-trial presentations of the CS alone were eliminated, as were five women and one man

[1] This study was carried out under the direction of Dr. K. W. Spence as part of a project conducted under contract N9onr-93802, Project 154-107 between the State University of Iowa and the Office of Naval Research.

TABLE 1

MEAN AND *SD* OF NUMBER OF CRs IN BLOCKS OF 20 TRIALS

Trials	Conditioning Group $N = 20$		Experimental Group $N = 20$	
	Mean	*SD*	Mean	*SD*
1–20	6.60	3.57	6.45	3.43
21–40	12.15	4.74	—	—
41–60	14.20	3.04	10.35	4.25

who indicated knowledge of the purpose of the experiment. One additional woman who exceeded the criterion used to exclude voluntary responders was dropped. This criterion, as in previous studies, was the occurrence of more than 50% CRs with a latency of less than 300 msec.

Apparatus and method of recording.—The equipment for recording eyelid closure and presenting the UCS was identical with that used in a previous study (**5**) and will not be described here. The CS was an increase in brightness of a 2¼-in.-diameter circular milk-glass disk from .05 apparent ft.-candles to 1.51 apparent ft.-candles. Duration of the light was 550 msec. The UCS was a 1.0-lb./sq.-in. air puff to the right eye, delivered through a .062-in.-diameter orifice. Duration of the UCS was limited to 50 msec. by a 110 v., 60-cycle AC-operated solenoid valve controlled by an electronic timer. The CS-UCS interval was 500 msec.

A CR was recorded whenever the record showed a deflection of 1 mm. or more in the interval of 200–500 msec. following the onset of the CS.

Procedure.—Each *S* was instructed to blink once to the ready signal (the word *ready*) and then to look at the disk, remaining as relaxed as possible and making no attempt to control his eye, until either the light went out or the air puff ceased. Following the reading of the instructions each *S* received three presentations of the CS alone and two presentations of the UCS alone in the following order: CS, CS, UCS, CS, UCS. All *S*s were then given 20 paired presentations of the CS and UCS and were assigned, following these trials, to the experimental or to the control (standard conditioning) group. The *S*s assigned to the experimental group then received 20 presentations of the puff alone followed by 20 paired presentations of the CS and UCS, while those in the control group received 40 paired-conditioning trials. Assignment of *S*s to the two groups was made so as to keep the frequency distributions of the number of CRs in the first 20 trials as similar as possible at all times. Thirty *S*s were assigned in this manner to each group. Intertrial intervals of 15, 20, and 25 sec., averaging 20 sec. and arranged by fixed schedule, were used throughout the 60 trials. The interval between the ready signal and the CS onset was 2, 3, or 4 sec., randomly varied. At the completion of the experimental session each *S* was questioned concerning his understanding of the purpose of the experiment and requested not to discuss the experiment with members of the class.

RESULTS

Performance curves for the two groups, based on the percentage of CRs occurring in successive blocks of trials for those *S*s who gave at least one CR in the first 20 trials, are presented in Fig. 1. Performance of the two groups is comparable over the first 20 conditioning trials. Contrary to the findings of Kimble, Mann, and Dufort, the conditioning group is superior to the experimental group over Trials 41–60, as seen in Table 1 and Fig. 1. The difference between the means of the two groups for this period, following differential treatment, provided a t value of 3.21, which for 39 df is significant beyond

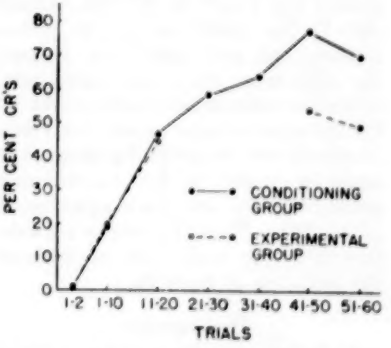

FIG. 1. Mean percentage of CRs for those 20 *S*s in each group who exhibited at least one CR during the first 20 trials. Points represent means over blocks of 10 trials, with the exception of the first point which represents performance on the first two trials.

the .01 level, utilizing a two-tailed test. A similar comparison was made between Trials 41–60 of the experimental group and Trials 21–40 of the conditioning group, at which points both groups had received the same number of CS-UCS pairings. Although the difference fails to reach significance (t of 1.3, which for 39 df has a $P > .20$), it again is in favor of the conditioning group.

In the data presented above only those 20 *Ss* in each group were included who exhibited at least one CR in the initial period. Inclusion of *Ss* who exhibited no CRs during the first 20 trials would obscure any possible adaptation effects of the UCS presented alone, for obviously those *Ss* could not be expected to give fewer CRs on their subsequent trials. However, excluding these *Ss* might preclude a fair comparison between representative groups of *Ss*. A further analysis was made including the 10 such *Ss* in each group who made no responses in the first 20 trials, increasing the total N in each group to 30. A Mann-Whitney U test of the difference between the two groups for Trials 41–60 was utilized due to the skewness of the distribution, and was significant beyond the .025 level for a two-tailed test. A similar comparison of Trials 41–60 for the experimental group and Trials 21–40 for the conditioning group was again in favor of the conditioning group, but did not reach significance ($P > .20$). Thus the results are substantially the same with or without the inclusion of these *Ss*.

Discussion

The finding that 20 interpolated UCS trials were not functionally equivalent to an equal number of paired trials is clearly at variance with the results of Kimble, Mann, and Dufort, and contrary to their theoretical prediction. They state, "The behavioral changes which are usually plotted as learning curves for eyelid conditioning probably reflect mainly the development of a second, *performance*, factor which is under the control of the UCS" (**2**, p. 415). This interpretation cannot account for the results of the present experiment, for trials with the puff alone brought no performance increment comparable to that produced by conditioning trials. Indeed, the level of performance of the experimental group during the final period of conditioning (Trials 41–60) was not as high as that of the control group during Trials 21–40. The experimental group at Trial 60 and the control group at Trial 40 had received an equal number of CS-UCS trials. Since the experimental group had received in addition 20 trials with the UCS alone, the poorer performance of this group is in line with the findings of Taylor (**6**) and MacDonald (**3**) revealing "adaptation" to the UCS with repeated presentations. Because some factor was increasing during the rise in the performance curves, the present findings support those formulations which postulate an associative factor which gradually increases during the course of conditioning (**1**, **4**).

Besides the small number of *Ss* used in the study by Kimble, Mann and Dufort, their procedure was different in a number of important respects from that used in the present experiment. It is common practice in conditioning experiments to allow for the possibility of temporal conditioning by varying the intertrial intervals. A ready signal is usually provided and the interval between it and the onset of the CS is also varied. Such procedures were not employed by Kimble, Mann and Dufort, who used a constant 15-sec. intertrial interval and apparently no ready signal. Since the CS occurred regularly every 15 sec., the possibility of temporal conditioning must be entertained. Such learning could take place during the trials with the UCS alone in the experimental group, increasing subsequent performance.

Another difference was their use of a CS which terminated 250 msec. prior to the onset of the UCS. This is in contrast with the utilization in the present experiment of a CS which terminated with the offset of the UCS. It is possible that the CS offset provided a cue to the approaching onset of the UCS and that avoidance responses were made at this time. Because of their latencies, such responses would not fall in the "voluntary" range (200–300 msec. after the CS onset), but would probably be recorded as CRs (300–500 msec.).

Work in this laboratory indicates that "voluntary responders" exhibit a characteristic pattern of response. Unfortunately, an all-or-none recording method was utilized by Kimble, Mann and Dufort, and no permanent record of the form of the eyelid closure was available. Voluntary responding might well have occurred quite apart from normal conditioning, and the inclusion in both groups of *S*s who gave a high incidence of such responses could have reduced the difference between the two groups on Trials 41–60.

While this discussion is only suggestive, the present results seem clear cut, and are accounted for by formulations which assume that increases in performance curves reflect a gradually increasing associative factor. Certainly the present results are at odds with both the findings and interpretation of Kimble, Mann and Dufort (2).

Summary

Two groups of 30 *S*s each were given 60 trials in an eyelid-conditioning experiment. The control group received 60 pairings of the CS and UCS in the normal conditioning procedure. The experimental group received only the UCS during Trials 21–40. Comparison of control and experimental groups on the last 20 trials revealed the experimental group to be significantly superior at the 1% level. Performance of the control group on Trials 21–40 was superior to that of the experimental group on Trials 41–60 but not significantly so. These results are interpreted as supporting the Hullian view that habit strength develops gradually during the experimental session as a function of CS-UCS pairings, and as failing to support a formulation which states that eyelid conditioning curves reflect primarily the growth of a motivational factor under the control of the UCS.

References

1. Hull, C. L. *Principles of behavior.* New York: Appleton-Century-Crofts, 1943.
2. Kimble, G. A., Mann, L. I., & Dufort, R. H. Classical and instrumental eyelid conditioning. *J. exp. Psychol.*, 1955, **49**, 407–417.
3. MacDonald, A. The effect of adaptation to the unconditioned stimulus upon the formation of conditioned avoidance responses. *J. exp. Psychol.*, 1946, **36**, 1–12.
4. Spence, K. W. *Behavior theory and conditioning.* New Haven: Yale Univer. Press, 1956.
5. Spence, K. W., & Taylor, J. A. Anxiety and strength of UCS as determiners of the amount of eyelid conditioning. *J. exp. Psychol.*, 1951, **42**, 183–188.
6. Taylor, J. A. Level of conditioning and intensity of the adaptation stimulus. *J. exp. Psychol.*, 1956, **51**, 127–130.

(Received for early publication October 17, 1956)

Journal of Experimental Psychology
Vol. 53, No. 3, 1957

ESTIMATES OF EXPECTED VALUE AS A FUNCTION OF DISTRIBUTION PARAMETERS[1]

ROBERT J. SCHREIBER

Dunlap and Associates, Inc., Stamford, Connecticut

In a recent series of studies, the author and others have attempted to derive a mathematical model which will predict an individual's responses in an accident-like situation. The qualitative differences between this type of situation and the standard learning situation are (*a*) the result of any behavioral response can range from very severe costs to slight rewards, and (*b*) the probability of experiencing a cost is very small. That is, each response could be said to be associated with a highly skewed probability density function of amount of reinforcement.

In an accident type of situation, the time between these responses is usually quite long and there are large numbers of interfering activities. An individual can be said to act as though he were estimating the parameters of the density functions, and the distortions due to time delays and interference can be said to result in misestimation. A mathematical model, then, must include the time delays, interference of the results of one response with another, and the amount that the individual contributes to the misestimation, independent of the two preceding influences.[2]

The goal of this study is to determine how the individual estimates the probability density function of results as a function of the density function itself, without interference of other responses and without time delay between events drawn from the distribution. Specifically, the purpose is to relate the first two moments (μ_W and σ_W^2) of the distribution of reward $[f(W)]$ to the first two moments (m_w and s^2_w) of the distribution of estimates $[g(w)]$.

The elimination of the two influences mentioned above defines the following constraints: (*a*) The *S* can choose only one course of action; (*b*) the *S* cannot forget the results of the previous choice of that action.

Procedure

The restrictions given above were satisfied as follows: each *S* was presented with four simultaneous events drawn from the parent distribution and was asked to estimate the next event. The individual may not be able to experience the four events simultaneously, but the time delay between the experiences is so small as to be negligible with respect to the ultimate purposes of the study.

An experimental trial consisted of showing *S* an index card on which were printed four numbers in a row, each number representing a sum of money up to $2.00 which ostensibly had to be drawn from a bowl. The *S* was to estimate the next amount to be drawn. An even number of numbers was chosen to prevent *S*s from using the middle value as an estimate, and four events were used to ensure that *S*'s visual span was not exceeded and that the task was not too simple. (These were confirmed by preliminary trials.)

For the parent distributions themselves, five means were used and each mean was associated

[1] This study was sponsored by the Commission on Accidental Trauma of the Armed Forces Epidemiological Board and was supported by the Office of the Surgeon General, Department of the Army, through Dunlap and Associates, Inc. The opinions expressed herein are the author's and are not necessarily those of the sponsoring agencies of the Department of Defense.

[2] The concepts and theory leading to these studies is reported in *Research on human risk behavior*, Stamford, Connecticut: Dunlap and Associates, Inc. Report submitted to the Armed Forces Epidemiological Board, 1954.

TABLE 1

NUMBERS USED AND THEIR DISTRIBUTION PARAMETERS

Mean	*SD*	Magnitudes			
$0.77	$0.08	.66	.74	.81	.87
.77	.12	.62	.70	.83	.93
.77	.16	.57	.67	.85	.99
.77	.20	.51	.66	.89	1.02
.77	.24	.52	.54	.94	1.06
.93	.08	.82	.90	.97	1.03
.93	.12	.78	.86	.99	1.09
.93	.16	.73	.83	1.01	1.15
.93	.20	.67	.82	1.05	1.18
.93	.24	.68	.70	1.10	1.22
1.09	.08	.97	1.06	1.13	1.19
1.09	.12	.94	1.02	1.15	1.25
1.09	.16	.89	.99	1.17	1.31
1.09	.20	.83	.98	1.21	1.34
1.09	.24	.84	.86	1.26	1.38
1.25	.08	1.14	1.22	1.29	1.35
1.25	.12	1.10	1.18	1.31	1.41
1.25	.16	1.05	1.15	1.33	1.47
1.25	.20	.99	1.14	1.37	1.50
1.25	.24	1.00	1.02	1.42	1.54
1.41	.08	1.30	1.38	1.45	1.51
1.41	.12	1.26	1.34	1.47	1.57
1.41	.16	1.21	1.31	1.49	1.63
1.41	.20	1.15	1.30	1.53	1.66
1.41	.24	1.16	1.18	1.58	1.70

with each of five different *SD*s. The actual events from each distribution were not chosen randomly because of the possibility of a highly skewed sample. Rather, the four numbers were chosen so that two would lie on each side of the mean and that the *SD*s (to two decimal places) would be as desired. Table 1 gives the means, *SD*s, and the actual numbers used. Each set of four numbers was arranged in all possible orders to eliminate the effect of order in the event that all the numbers were not simultaneously experienced. Each ordering was typed on a 3 × 5-in. index card and coded on the back to indicate the distribution from which it had been drawn. The 600 cards (25 distributions × 24 orderings) were then randomized according to a table of random numbers and divided into decks of 30, which were presented to each *S* in the same order. Each *S*, then, received 24 replications of a 5 × 5 factorial design (means × *SD*s).[3] Five *S*s were used, all male college graduates with no knowledge of the purpose of the experiment.

The cards were shown to *S* through an aperture for 3 sec. each. The *S*s were instructed that these amounts of money had been drawn from a bowl for them and that they were to estimate the next sum that would be drawn. They were to give this estimate before the card was withdrawn, and after a few training sessions this was done with apparent ease. Approximately 5 sec. were allowed between exposures. At the end of each deck of 30 cards, 2 min. of rest was allowed, and after each four decks (120 cards), a 10-min. rest period was given. The instructions were re-read to *S* before each set of four decks.

[3] All *S*s were questioned about the replication, and none had been aware that there were only 25 different sets of numbers.

RESULTS

The 24 estimates of each set were plotted and examined for each *S*. Although some of the 120 distributions showed marked irregularities, the irregularities were not sufficiently regular in any way to justify any conclusion. A random sample of 15 of the distributions were fitted to the normal distribution, and the fit was visibly good; chi-square tests did not permit rejection of the normal distribution.

The means and *SD*s of the estimates (m_w and s_w) were compared with the means and *SD*s (μ_W and σ_W) of the original numbers for all individuals as a population.

μ_W vs. m_w.—Table 2 shows the mean estimates as a function of both μ_W and σ_W. Comparison of the numbers in the body of the table with the column headings indicates remarkably good agreement. Without calculation, it can be seen that, with the *SD*s given in Table 3, none of the

TABLE 2

MEAN ESTIMATES AS A FUNCTION OF BOTH μ_W AND σ_W

σ_W	μ_W				
	.77	.93	1.09	1.25	1.41
.08	.771	.937	1.095	1.252	1.406
.12	.781	.923	1.110	1.250	1.387
.16	.773	.942	1.116	1.246	1.404
.20	.783	.946	1.109	1.253	1.388
.24	.770	.954	1.086	1.251	1.384

TABLE 3

*SD*s of Estimates

σ_W	μ_W				
	.77	.93	1.09	1.25	1.41
.08	.028	.022	.028	.027	.039
.12	.029	.033	.037	.032	.042
.16	.038	.038	.058	.033	.066
.20	.049	.047	.047	.035	.053
.24	.043	.070	.044	.049	.044

mean estimates is more than about $\frac{1}{2}$ *SD* from the μ_W. Thus, the differences are not significant at the 5% level. The mean estimates of the next drawing from a distribution whose mean is $1.41 are all lower than $1.41. However, for the matrix as a whole, there is no significant difference between the proportions of under- and overestimations.

σ_W *vs.* S_w.—Table 3 shows both of these. In order to test the hypothesis that the S_w are equal for all σ_W, the experimental data were reformulated as a 5 × 5 × 5 factorial design (means × *SD*s × Ss) with the *SD*s of the estimates as the measurements. The analysis of variance resulted in Table 4. The differences in S_w as a function of σ_W found in the analysis of variance seem to approach a relationship where S_w increases as σ_W increases.

TABLE 4

Analysis of Variance

Source	*df*	Mean Square	*F*	*P*
Means	4	$.53 \times 10^{-2}$	2.44	>.05
*SD*s	4	1.57×10^{-2}	7.35	<.01
*S*s	4	2.78×10^{-2}	12.80	<.01
Error	112	$.22 \times 10^{-2}$		

S_w *vs.* μ_W.—Comparing the cell entries with the column headings in Table 3 indicates a slight tendency for the S_w to increase as μ_W increases. However, this increase is not sufficiently regular to be able to draw any conclusions with confidence.

m_w *vs.* σ_W.—Comparing the cell entries of Table 2 with the row headings provides a comparison of m_w and σ_W. There appears to be no relationship—indeed, if one does exist, it is so small as to be negligible.

Summary

Subjects were asked to estimate the next amount of money they would draw from a bowl given four previous drawings. The results indicate that (*a*) the estimates cannot be differentiated from a normal distribution, (*b*) the mean of the estimates seems to be an unbiased estimator of the mean of the distribution from which the drawings are made, and (*c*) the dispersion of the estimates increases with the dispersion of the four previous drawings.

(Received for early publication
August 27, 1956)